# FINAL TECHNICAL REPORT

# M-071 CRITICAL DATA ANALYSIS

PRINCIPAL INVESTIGATOR: D. M. HEGSTED

PERIOD COVERED BY REPORT:

JUNE 27, 1973-DECEMBER 31,1975

GRANT NO. NAS 9-13370

ADDRESS OF GRANTEE INSTITUTION:

HARVARD UNIVERSITY
1350 MASSACHUSETTS AVENUE
CAMBRIDGE, MASSACHUSETTS

## INTRODUCTION

A prototype balance study was conducted on earth [SMEAT] prior to the balance studies conducted in Skylab itself. The data collected were the daily dietary intake of 6 minerals and nitrogen, and the fecal and urinary outputs on each of three astronauts. Our goals have been to establish methods of analysis that would allow computing the net balances along with the error bounds of these estimates; to search for structure in the data which might allow a tightening of these error bounds and to possibly improve the design of future balance studies; to evaluate the biological implications of the SMEAT data themselves and relative to data which will be obtained in Skylab or other metabolic balance studies.

## SUMMARY OF CONCLUSIONS

We have set out the essential statistical issues, showing what quantities need to be estimated and establishing the scope of the inference associated with alternative variance estimates. The procedures for obtaining the final variability due both to errors of measurement and "total error" (total = measurement and biological variability) are exhibited. The previous document by Arnold [1] dealt only with the former, very restricted, issue.

Quantitative estimates of both the measurement and total variance in the net average apparent retention as a function of the duration of the study have been obtained. These have implications in the design of future balance studies. The mean net retention and its variance establish

the reproducibility of the balance study for each of the substances considered, leading to consistent results across subjects. We establish that no inter-dependencies or periodicities allow significant reductions of the above variance estimates.

The "biological significance" of the net retentions which have been observed is considered. We conclude, unfortunately, that the interpretation of balance studies of this kind present unresolvable difficulties and it is uncertain whether this study [SMEAT] can serve as a base line for the interpretation of the data from Skylab.

## SECTION I

The Data Base. We have concentrated on the SMEAT data available for nitrogen (N), calcium (Ca) and potassium (K). The principle variables in the analysis are, for each of 100 days, the total content of the daily diet and the urine and feces. Two features of the data--[3 astronauts × 3 minerals × 1 input + 2 excretory modes] = 27 data series--that conditioned our investigation are: the diets had an approximate (but not equal) 6 day periodicity with only small variations among the 6 days and the fecal series had highly fluctuating values with zero output on many days. [The series also contained a few irregularities like missing values, duplicate readings, etc. Individual judgments were made regarding the most appropriate single value to replace each of these irregularities; using the sample information efficiently. There were not many such problems.]

Definitions and Statistical Properties of the Fundamental Balance Quantities. If $I_i$, $U_i$, and $F_i$ represent the mineral content of a particular mineral (in g, say) in the $i^{th}$ day's diet, urine and feces, respectively,

then we may define the net retention on day i as

(1) $$NET_i = I_i - U_i - F_i$$

and the average cumulative retention through day k as

(2) $$CUM_k = \frac{\sum_{i=1}^{k} NET_i}{k}$$

Regarding $I_i$, $U_i$ and $F_i$ as random variables possessing a mean (E) and a variance ($\sigma^2$), the same for all i, we find by the rules of elementary statistics that

(3) $$E(CUM_k) = \frac{\sum_{i=1}^{k} E(NET_i)}{k} = \frac{k \cdot E(NET)}{k} = E(NET)$$

and similarly

(4) $$\sigma^2(CUM_k) = \frac{k\,\sigma^2(NET)}{k^2} = \frac{\sigma^2(NET)}{k}$$

Also notice that the right-hand sides are given by

(5) $$E(NET) = E(I) - E(U) - E(F)$$

and

(6) $$\sigma^2(NET) = \sigma^2(I) + \sigma^2(U) + \sigma^2(F).$$

We suppress the subscript i wherever possible above by the assumption that E and $\sigma^2$ of $NET_i$ are the same for all i, equal respectively to E(NET) and $\sigma^2(NET)$.

In a balance study of duration k days, the principal quantities whose estimates we desire are

$$E(CUM_k) \quad \text{and} \quad \sigma^2(CUM_k)$$

We may obtain confidence intervals for estimates of the average retention E(NET) as a function of MBE duration as, in the conventional way,

(7) $\hat{E}(CUM_k) \pm 2\sqrt{\hat{\sigma}^2(CUM_k)}$

where the "hats" denote data-based estimates of the corresponding true parameters.

Equations (3) through (7) demonstrate, then, that to establish error bounds on the principal quantity of interest, E(NET), is effectively a question of finding estimates of

$$\sigma^2(I),\ \sigma^2(U),\ \sigma^2(F),$$

the variances of the mineral contents in one day's diet--urine and feces.

The SMEAT experience consisted of "pre-, in- and post-chamber phases" and later discussion in fact demonstrates that it is not always appropriate to assume that the above quantities $\sigma^2$ are independent of the day-number i. But in the present treatment we find this assumption tenable so long as i is not too near 1 or 100, i.e., so long as we eliminate the highly variable beginning and end of the experiment.

The issue of the independence of E on i has also to do with the important issue of <u>trend in net retention over time</u> which we will return to later.

<u>Scope of Desired Inference Determines Appropriate Estimates of Variance--Status of the Arnold Document [1].</u> We may regard the data series for I, U, and F in essentially two distinct ways; the appropriate estimates

$$\hat{\sigma}^2(I),\ \hat{\sigma}^2(U),\ \hat{\sigma}^2(F)$$

will be correspondingly different and the scopes of the corresponding inferences will be also.

The Restricted Inference. This corresponds to answering a question that may be put this way: "What is the estimate of this astronaut's net retention in this particular metabolic balance experiment (MBE)?" If all quantities were measured with perfect precision (and if no "edge effects" occurred because of where the starting and the stopping of the MBE "caught" the I, U, and F series), then this question could be answered with perfect precision, i.e., the confidence interval (7) would consist of the single point

$$\hat{E}(CUM_D) = \frac{\sum_{i=1}^{D} NET_i}{D}$$

where

D = number of days in MBE

Status of the Arnold Document [1]. Absent perfect precision (but still ignoring edge effects), the width of (7) is determined by estimates of the errors of measurement in $I_i$, $U_i$, and $F_i$.

This is the way Arnold [1] regarded the NASA MBE's; he set forth a model that was intended to allow obtaining estimates of

$$\sigma^2(I),\ \sigma^2(U),\ \sigma^2(F)$$

where $\sigma^2$ now denotes measurement error only. He then asserted that this is the only source of variability that need be taken into account.

The present study regards the Arnold document as only relevant to the narrow question stated above and asserts that this is in fact only a rather specialized and peripheral issue. This section sets forth what we take to be the central question at issue. Appendix II, however, does build upon the model of the Arnold document, changing it somewhat and using the data themselves (the U and F series as well as others described in Appendix II)

to compute estimates of the measurement uncertainty, because we do regard as important the issue of what proportion of the variability relevant to our inferences is in fact due to measurement error. Our result is that this proportion is small, indeed usually negligible in all cases.

The Appropriate Inference. The question associated with the relevant inference can be posed in this way: "If we regard the present MBE as a particular realization of an MBE under these experimental conditions, what inferences can we make about the uncertainties in net retention under these experimental conditions?" In the context of the Skylab experiment, the difference is between asking "What happened in this experiment to this astronaut while in flight?" and asking "What is likely to happen to this astronaut if he were to repeat the experiment?" [Since SMEAT is regarded as a prototype of Skylab, it is relevant to ask what inferences can be drawn from a comparison of the data obtained in SMEAT and that obtained in Skylab? This cannot be answered without data from Skylab but this is discussed below.]

To make inferences with this wider scope, we must contend not only with measurement errors but also with the "biological" variability that would be present even if all measurements could be made with perfect precision. Indeed, we find that to make inferences with this wider scope requires paying a high price in the efficiency of our estimates--i.e., our confidence interval (7) that makes a statement about the random process from which we have observed a single realization proves to be much wider than that pertaining only to the particular realization. This follows from the

unfortunate fact that biological variability far outweighs measurement error. The conceptual distinction reflects itself dramatically in the very different estimates of $\sigma^2(I)$, $\sigma^2(U)$, and $\sigma^2(F)$ that are to be substituted in equations (4), (6) and (7).

Summary of Estimates. Table 1 provides our estimates of the mean and the variance of I, U, F, and NET for each subject where

$$NET = I - U - F$$

The three subjects are identified as S, C, and P (or as SPT, the scientist-pilot; CDR, the commander-copilot and PLT, pilot). One value for $\hat{E}$ and two values for $\hat{\sigma}^2$ are included, the first relates to the entire study and the second roughly to the in-chamber period. The in-chamber period was the most orderly phase of SMEAT and the estimated $\hat{\sigma}^2$ are usually, but not always, somewhat smaller than for the entire period. The second $\hat{\sigma}^2$ thus presumably represent the minimal variance that can be expected in such studies.

Section II describes how the estimates of $\hat{\sigma}^2$ were computed. Appendix I documents our efforts to reduce these large variances by searching for patterns in the data. Our principal finding is that, unfortunately, substantial reductions in variance are not available and the variances presented in Table I are realistic estimates of the variability inherent in metabolic balance studies.

Table 2 summarizes the findings derived in Appendix II relative to the fraction of the $\hat{\sigma}^2$ which is due to errors of measurement. The total variance has been decomposed into the variance due to biologic and measurement components;

$$\sigma^2_{\text{total}} = \sigma^2_{\text{biological}} + \sigma^2_{\text{measurement}}$$

Table 1. $\hat{E}$, $\hat{\sigma}^2$ for I, U, F and NET (g and $g^2$, respectively)

| | | | I(input) | U(urine) | F(feces) | NET |
|---|---|---|---|---|---|---|
| "S" | | | | | | |
| Nitrogen | $\hat{E}$ | | 16.819 | 15.202 | 0.9367 | 0.680 |
| | $\hat{\sigma}^2$ | (diet periods 1-12) | 2.1962 | 36.040 | 0.4166 | 38.653 |
| | $\hat{\sigma}^2$ | (diet periods 4-12) | 0.3590 | 29.521 | 0.4620 | 30.342 |
| Calcium | $\hat{E}$ | | 0.8593 | 0.1172 | 0.5695 | 0.1725 |
| | $\hat{\sigma}^2$ | (diet periods 1-16) | 0.001510 | 0.0007039 | 0.1951 | 0.1973 |
| | $\hat{\sigma}^2$ | (diet periods 4-13) | 0.0005960 | 0.0006063 | 0.1367 | 0.1379 |
| Potassium | $\hat{E}$ | | 4.0708 | 2.8674 | 0.3303 | 0.8732 |
| | $\hat{\sigma}^2$ | (diet periods 1-16) | 0.5930 | 0.5087 | 0.08932 | 0.6573 |
| | $\hat{\sigma}^2$ | (diet periods 4-13) | 0.1107 | 0.4248 | 0.08461 | 0.5205 |
| "C" | | | | | | |
| Nitrogen | $\hat{E}$ | | 16.1812 | 13.296 | 0.9205 | 1.9675 |
| | $\hat{\sigma}^2$ | (diet periods 1-16) | 0.2798 | 10.481 | 1.1496 | 11.910 |
| | $\hat{\sigma}^2$ | (diet periods 4-13) | 0.2371 | 8.7675 | 0.9663 | 9.971 |
| Calcium | $\hat{E}$ | | 0.8498 | 0.2466 | 0.4418 | 0.1619 |
| | $\hat{\sigma}^2$ | (diet periods 1-16) | 0.0001327 | 0.001744 | 0.2726 | 0.2745 |
| | $\hat{\sigma}^2$ | (diet periods 4-13) | 0.00002188 | 0.001239 | 0.2779 | 0.2791 |
| Potassium | $\hat{E}$ | | 3.8781 | 3.05947 | 0.2581 | 0.5605 |
| | $\hat{\sigma}^2$ | (diet periods 1-16) | 0.007779 | 0.30776 | 0.9127 | 0.4068 |
| | $\hat{\sigma}^2$ | (diet periods 4-13) | 0.001291 | 0.21774 | 0.0933 | 0.3123 |
| "P" | | | | | | |
| Nitrogen | $\hat{E}$ | | 16.798 | 13.8117 | 1.1727 | 1.8136 |
| | $\hat{\sigma}^2$ | (diet periods 1-12) | 0.5450 | 13.9891 | 0.4826 | 15.0167 |
| | $\hat{\sigma}^2$ | (diet periods 4-12) | 0.5450 | 13.1447 | 0.4284 | 14.1182 |
| Calcium | $\hat{E}$ | | 0.8484 | 0.1223 | 0.5765 | 0.1496 |
| | $\hat{\sigma}^2$ | (diet periods 1-16) | 0.0001494 | 0.0004431 | 0.1492 | 0.1498 |
| | $\hat{\sigma}^2$ | (diet periods 4-13) | 0.00002137 | 0.0003564 | 0.1207 | 0.1211 |
| Potassium | $\hat{E}$ | | 3.9701 | 3.016 | 0.3241 | 0.6308 |
| | $\hat{\sigma}^2$ | (diet periods 1-16) | 0.009418 | 0.1802 | 0.7413 | 0.9309 |
| | $\hat{\sigma}^2$ | (diet periods 4-13) | 0.003722 | 0.2050 | 0.04208 | 0.2508 |

and in Table 3

$$\frac{100\ \hat{\sigma}^2_{\text{measurement}}}{\hat{\sigma}^2_{\text{total}}}$$

and is shown for U, F, and NET of Ca and N for astronaut C.

It is apparent that measurement error contributes only minimally to the uncertainty associated with the mean values observed in the study. [N.B. As section II will make clear, our estimates of $\hat{\sigma}^2$ for I in Table 1 do not include any variance contribution from the purely cyclical part of the diet variability as, indeed, one would not wish them to. Also, notice that any deviation from the pure cyclicality must, in fact, be due to measurement error and/or lapses in the experimental procedure. Absent the latter, then, the entries for I in Table 3 must be 100% and any deviation therefrom must be due either to accidents peculiar to this experiment or to artifacts of the calculation procedure--neither of which would be important to document. Indeed, these causes reflect themselves in $\hat{\sigma}^2_{\text{measurement}}(\text{I})$ for Ca for astronaut C in Table 2 being greater than both $\hat{\sigma}^2(\text{I})$'s in Table 1. This is a nonsensical result if taken literally but Table 2 was calculated from day 2 of C's diet.]

The fact that the percentages in Table 3 are so small demonstrates that little can be gained by improvements in analytical accuracy. We have accordingly devoted considerable effort to exploring whether the biologic variation can be reduced by exploiting interdependencies among the data (Appendix I). Unfortunately, no means of reducing the variance has been found and the values presented must be considered to be representative of the information gained from this kind of metabolic balance study.

Table 2. Measurement-Variance Estimates, $g^2$ (CDR; day 2 diet; "typical" day's output)

| | Input | Urine | Feces | NET |
|---|---|---|---|---|
| Nitrogen | 0.046178 | 0.1371 | 0.01162609 | 0.1950 |
| Calcium | 0.00016909 | 0.00008968 | 0.00022533 | 0.0004850 |

Table 3. Percent of Total-Variance Estimate (for periods 4-13) Due to Measurement Errors (CDR)

| | Urine | Feces | NET |
|---|---|---|---|
| Nitrogen | 1.6 | 1.2 | 2.0 |
| Calcium | 0.7 | 0.08 | 0.17 |

Biologic Implication of the SMEAT Data. It would seem certain that in well-nourished young men who consume a "normal" diet and who do not change weight to a significant degree that the body composition must be approximately constant over rather long periods of time. If this is true, the amount of various nutrients lost from the body must be approximately the same as the amounts consumed. This is the fundamental assumption upon which balance studies such as SMEAT and Skylab are based.

Consider now the mean retentions for the various nutrients shown in Table 1. For astronaut "S" the average retention of nitrogen is 0.68 g/day. Thus, over an 100 day period one would calculate that a total of 68 g of nitrogen would be retained if this value is correct. This would be equivalent to 425 g of protein (protein = N × 6.25) or 2360 g of new tissue if one accepts the conventional figure of 18% of protein in average tissue. The latter assumption would also lead to an estimate of approximately 12.6 kg of protein in an average man weighing 70 kg. Thus, the net retention of nitrogen would lead to an estimate that the amount of nitrogen retained during an 100 day study would increase the total body protein approximately 3.4%,

$$\left(\frac{0.425 \times 100}{12.6} = 3.37\right)$$

A similar calculation for astronaut "C," showing a mean retention of 1.96 g N/day leads to an estimated increase in the body protein of 1.225 kg in 100 days or approximately an 9.7% increase in the total body protein. Since in the nitrogen retention data there is no apparent tendency for the retention to decrease with time (Section II, Figs. CUM1 - CUM3), these values appear to be representative of these subjects under these conditions. Thus, had the study

been continued for a year the net observed retentions would have been approximately 3.5 times the values calculated above and clearly led to a nonsensical result.

For purposes of this discussion the results of some of the fairly recent nitrogen balance studies have been plotted in Fig. 1.

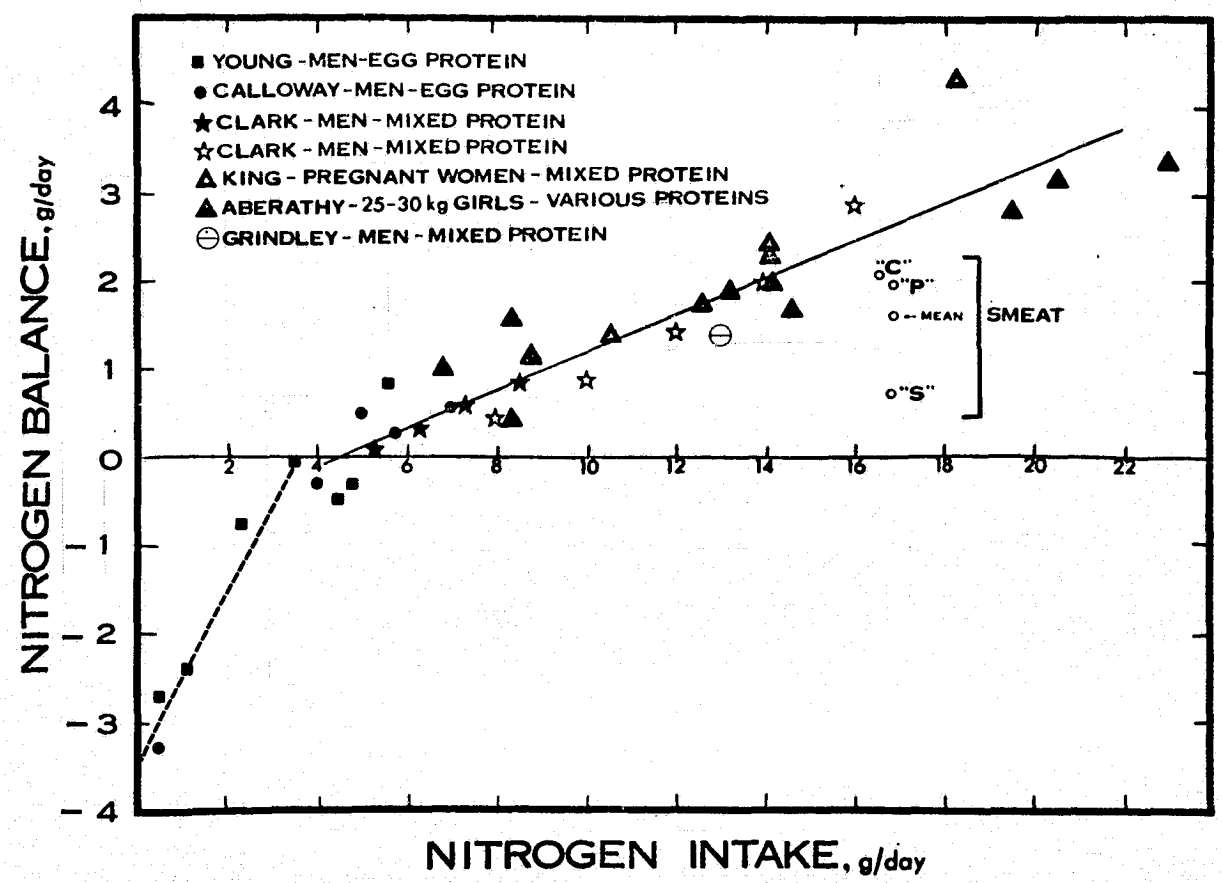



Fig. 1

In the studies of Young et al. [2] and Calloway and Margen [3], egg protein was fed to young adult men at low levels. In the two papers of Clark et al. [4,5] young men were fed a mixed diet which supplied from 5 to 16 g of nitrogen/person/day. The data of King et al. [6] were obtained with 10 young

pregnant women who received up to 18 g N/day. These women were 4 to 7 years postmenarche and were considered to be mature. Thus, the nitrogen retention would be expected to be that required for the formation of the fetus and the accessory tissues related to pregnancy. Abernathy et al. [7] studied young girls weighing from 25-30 kg. Standard tables indicate that such girls should be gaining about 6-7 g/day. The results were reported as mg N retained/kg/day and these values have been multiplied by 65 to make the results more or less comparable to the other data obtained with adults. The values plotted would represent the retentions expected in adults growing at the same rate as the children. No other corrections for body size were made in the data plotted although there were substantial differences in the weight of the subjects studied. Finally, the single average value obtained by Grindley [8] is shown. This was obtained with 23 young men who were continuously investigated over a 220 day period. They received a more or less normal diet which supplied about 13 g N/day.

The regression line in the upper right hand portion of the figure was drawn by inspection. Since each point represents variable numbers of subjects, the true regression line would be obtained by appropriate weighting for the number of subjects. It is not likely, however, that this would deviate greatly from the line shown.

It is quite remarkable that in this heterogeneous group of experiments in which the subjects differed in age, sex, size and physiological state, and in which the experimental conditions varied as well as the nature of the protein fed, nitrogen intake is the only variable which appears to explain the nitrogen balance. Indeed, the variability around this line when the intakes exceed

5 g N/day does not appear to be much greater than that observed in the individual experiments. There is no indication that a "plateau" is reached as the intake is raised.

The average retentions observed for each of the subjects in SMEAT as well as the grand mean for the three subjects are also shown in Fig. 1. It is apparent that although most of the balance studies reported in the literature are for relatively short periods, usually two weeks or so on the diet and measurements being made during the latter part of the study, there is little indication that the data from SMEAT vary appreciably from the general data obtained. Since neither SMEAT nor the long-term balance study of Grindley depart significantly from the general findings, "adaptation" to the diet does not appear to be a significant factor in the results obtained. This also confirms the conclusion (Appendix I) that in nitrogen balance studies little is gained by extension of the study period.

The values above the balance line indicate a "requirement" (where the balance is zero and excretion approximates intake) of about 0.5 g N/person/day. This would be equivalent to about 31 g of protein/day for a 65 kg adult. Similar estimates have been reached by many investigators in the past.

The slope of the regression line at the higher intakes indicates a retention of approximately 20% of the intake above 5 g/day. That is, there is a retention of approximately 2 g N/day at an intake of 14 g/day and $\frac{2}{14-5} \times 100 = \pm 20\%$. A retention of 2 g of nitrogen is equivalent to 12.5 g of protein or about 70 g of tissue if one accepts the conventional figure of 18% protein in body tissue. This would result in a gain of over 25 kg/year and clearly cannot be typical of young men who do, in

fact, often consume diets containing 14 g of nitrogen or more.

Since nitrogen and other nutrients are lost through the skin as sweat, desquamation of skin, hair, nails, etc., and these are not measured in the usual balance study, the positive balances are falsely high. Attempts to measure such losses, however, indicate that under ordinary conditions they are of the order of 0.3 to 0.5 g N/day [9,10]. If one accepts the higher value and subtracts 0.5 from each value, the line is shifted to the right but the conclusions are not changed to any substantial degree.

Forbes [11] has suggested that the balances are erroneous because modification of body composition after a dietary change may occur slowly and exponentially with time. The usual short balance periods of two weeks or so would be too short to identify stable conditions. There is probably sufficient data to show that this suggestion is not valid when the protein intake is lowered and that relatively stable conditions are found within a week or 10 days. Adequate data after the dietary protein is raised have not been found. However, the data reported by Grindley and the present findings of SMEAT lend no support to the argument presented by Forbes, at least insofar as nitrogen balance data are concerned.

Wallace [12] has presented the reasoned argument that there is a consistent "bias" in balance studies which inevitably lead to falsely high apparent retentions. It can be assumed that in most balance trials the intake will be overestimated since the subject may not eat all of the food offered but has little or no opportunity to consume more than offered. Excretion will generally be underestimated since some loss is probably inevitable but it is difficult to collect more excreta than are actually produced. In the context of SMEAT when the men were actually

in the chamber, we see little possibility that the high net retentions could be explained on this basis. The average net retention of the 3 subjects was 1.48 g N/day, approximately 13% of the nitrogen intake above 5 g/day. Retentions this high would require overestimate of the intake by about 6 to 7 % and similar underestimates of the excretion or larger errors in one or the other. This seems most improbable.

The average net retention of potassium was 0.688 g/man/day or some 68 g/100 days. Considering the fact that the estimated total body potassium is of the order of 120 g, it would follow that if this were a measure of true change in body potassium, the body potassium would double over a 200 day period. It can be assumed that relatively large amounts of potassium are lost through the skin and the results of potassium balance studies are not meaningful.

The calcium balance data are more difficult to interpret. Although they show a consistent and improbable retention (average retentions = 161 mg/man/day or a total net retention over 100 days of 16 g of calcium), the amount retained would be relatively small compared to the total body calcium content. If it be assumed that the total body contains something of the order of 1.5% calcium, the total body content would be of the order of 1.05 kg. A net retention of 16 g during a 100 day period would amount to only 1.5% of the total body calcium, apparently within the range of possible retention without observable physiological consequences. Of course, if this rate of retentions is extended to long periods, it leads to ridiculous conclusions. The cumulative retention curves (Section II, Figs. CUM4 - CUM6), however, indicate a gradual reduction in the degree of calcium retention suggesting, possibly, an adapatation to the diet fed.

Duncan [13] reviewed the results of calcium and phosphorus balance trials in ruminants. Carcass composition is better known in animals than in man and in some of the trials animals were killed for carcass analysis. Her conclusions, in part, were that "cumulative estimates of retention add up to totals far in excess of likely increments" and that "Careful examination of published metabolism trials has revealed no probable explanation. The discrepancy cannot arise from losses through the skin and is not likely due to failure to collect all of the excreta." These results are important relative to the nitrogen balance studies since there has been speculation that the false positive nitrogen retentions might be explained by reduction and loss of nitrogen as gaseous $N_2$ or ammonia. Volatile losses clearly cannot explain false retentions of calcium.

If SMEAT is considered to be a typical metabolic balance study or a prototype of Skylab, the results of Skylab will presumably be compared with the data obtained in SMEAT. The confidence intervals of the NET of the various nutrients are, therefore, important. Table 4 from Appendix I presents the most optimistic estimates of the confidence intervals for each nutrient for each subject.

Table 4. Long-Term Confidence Intervals for Net Retention (in g)

| Nutrient | Subjects | | |
|---|---|---|---|
| | S | C | P |
| Nitrogen | 0.60-1.96 | 1.30-2.60 | 1.00-2.70 |
| Calcium | 0.10-0.25 | 0.06-0.27 | 0.08-0.22 |
| Potassium | 0.73-1.02 | 0.45-0.67 | 0.53-0.73 |

Considering the large variance attached to NET of each subject and the rather large differences in mean NET and variance between subjects, the difficulties or limitations in comparing one study with another with different subjects and under different environmental conditions becomes obvious.

From the net retentions of nitrogen and potassium, it is clear that the findings are unrealistic and it may be assumed that the calcium balances are also unrealistically high. Although it cannot be proven, the only possible explanation that seems reasonable is that other losses (sweat, skin, hair, etc.) which are not measured are higher than the conventional estimates. Very large losses of certain nutrients--calcium, for example--have been reported in men under high environmental temperatures [14]. The losses reported, in fact, were so high that they could not possibly be representative of most people living in hot climates. Nevertheless, if this is the explanation of the results obtained, both of the average retention and the differences observed between experimental subjects, it cannot be assumed that meaningful results will be obtained by applying a standard correction for such losses as is often assumed. They will have to be measured if, as the data suggest, they may be the primary cause of differences and unexplained results obtained with balance studies. This is probably particularly true in Skylab where the effects of the environment on dermal losses are completely unexplored.

The difficulties associated with accurate measures of skin losses are also well known. The dangers of contamination are great; collecting skin losses requires large volumes of fluid and imposes analytical difficulties; cooperation from subjects is also difficult to maintain. It seems unlikely that meaningful experiments will be done except under unusual conditions.

There is great need, therefore, for independent measures of body composition. In the studies of King et al. [6], the average values of which are plotted in Fig. 1, the changes in total body $^{40}K$ were determined. Assuming the conventional relationships between body K and body N, the $^{40}K$ counts yielded estimates of nitrogen retention ranging from 29 to 127% of the values estimated from the nitrogen balance studies with an average value of 67%. It appears from the published data that there was no correlation between the estimates derived from the balance studies and those from the $^{40}K$ counts. The mean value of 67% in these pregnant women who were obviously retaining nitrogen appears to be more realistic than the nitrogen balance data.

We unfortunately must conclude that however accurately metabolic balance studies are conducted they will inevitably lead to uncertain results. The usual experience is that the more nutrient fed the more "apparent retention" is obtained and that the results are not made more interpretable by very long-term studies. Since the reasons for the "errors" are unknown (even though skin losses seem the most logical explanation), comparative studies with different environmental conditions or different subjects must be viewed with great caution and suspicion. Calcium balances are particularly difficult to interpret. Adaptation to new conditions may require months and the losses or gains in total body calcium which appear to occur, as in SMEAT, may be insignificant in terms of the total body supply. Clearly, one cannot project losses observed over even a few months to long periods.

Studies utilizing activation analysis for estimation of total body calcium or nitrogen or the measurement of total body potassium by $^{40}K$ counts or the estimation of total potassium pools appear to offer the best direction for research. Even though the accuracy of such measures is difficult

to establish, they should be satisfactory for comparative measurements, before and after experimental treatments, as in the exposure to weightless conditions. Until such independent measures of total body composition are available, the interpretation of balance studies will remain questionable.

## SECTION II

Our program in the remainder of the text and Appendix I shall be (1) to explain how the estimates of total variances (and means) in Table 1 were obtained; (2) to assess their implications regarding MBE accuracy and design; (3) to explore where the larger variance-contributions are and whether they can be reduced significantly.

Appendix I essentially shows that these variance estimates in Table 1 represent realistic lower limits to the variabilities inherent in the MBE procedures for each of the three substances--N, Ca and K.

Derivation of Table 1. Does diet "drive" excretion? One question in our search for structure to the basic data series, I, U and F, is: to what extent are the excretory variables driven by the periodic diet? Because the diet is, with only moderate departures, repeated every 6 days, i.e., there are 6 one-day menus for each of the three astronauts that are administered cyclically, this question centers around searching for periodicity in U and F with a 6-day cycle. Appendix I applies to these series several statistical techniques well suited to examining cyclical features of time series; these include autocorrelations and autoregressions, and spectral and cross-spectral analyses.

We should like our summary estimates of total variability not to reflect any variation in I, U and F due to these 6-day periodicities. In order to obtain even more "optimistic" variance estimates, we should like to regard as being "known in advance" any trends across time between 6-day periods in these series.

Interpretation of Variance Estimates. Each $\hat{\sigma}^2$ in Table 1 is accordingly the residual mean square of a two-way analysis of variance (AV) of each of the I, U and F series viewed as having 6-day periods, i.e., from each series $\{U_i\}$ say, we define a matrix $U$ whose $(l,m)$ element obeys the relation

(8) $$u_{lm} = U_{6l+m}$$

This amounts to regarding each such $u_{lm}$ as

(9) $$u_{lm} = E(U) + \tau_l + d_m + \varepsilon_{lm}$$

where $E(U)$ is an overall mean, $\tau_l$ is a "period-effect" (which would increase with increasing $l$, say, if the astronaut tended to lose more mineral in the urine as the MBE progressed), $d_m$ is a "diet-day effect" (which might reflect the 6-day sequence of the diet "loadings" of the mineral), and $\varepsilon_{lm}$ is a random error, identically distributed for all $(l,m)$. The residual mean square is an estimate of $\sigma^2(\varepsilon)$ and is thus an optimistic estimate of $\sigma^2(U)$ as it is less than the total mean square

$$[\sum_l \sum_m (U_{lm} - \hat{E}(U))^2/(\text{length of series})]$$

by an amount that increases as the rows and columns of $\underset{\sim}{U}$ do display the effects allowed for in increasingly systematic degrees. Because any departure from randomness between 6-day periods (that would give a net higher variance than under-randomness) would reduce the mean sum of squares, the variance entries in Table 1 are optimistic in that they reflect a reduction due not necessarily (or only) to trend but rather to a wider class of alternatives to the assumption of between-period randomness.

Contributions to Total Variance $\hat{\sigma}^2$. Table 1 shows that, for N and K, the primary source of the variability in NET is the urine, i.e.,

$$\hat{\sigma}^2(U) >> \hat{\sigma}^2(F),$$

for calcium, by contrast,

$$\hat{\sigma}^2(F) >> \hat{\sigma}^2(U).$$

The latter fact reflects the violently fluctuating character of all the fecal series and the fact that more calcium is excreted in the feces than in the urine; the former reflects the fact that more N and K are excreted in the urine than in the feces. Figs. C1-C6 are the graphs of I, U and F for N and for Ca respectively for astronaut C.

Notation Convention. From now on we refer to data series in the forms indicated in the titles of the figures as: letter-name-substance where letter = astronaut identifying letter (S,C or P), name = the variable name (I, U, F, NET, etc.); substance = substance (N, Ca or K). Thus SNET Ca refers to the net retention of Ca in subject S.

Implications of Table 1. As discussed previously, we note that all values of $\hat{E}$(NET) are positive and these values cannot represent reasonable estimates of actual retention. The retentions are clearly artificially high since dermal excretion was not measured. Estimates available currently for the amount of dermal losses do not appear to be large enough to account for these retentions, yet this provides the only reasonable explanation of the results obtained. More important, however, would be the implication that if all three of the astronauts were, in fact, "in balance" then the dermal losses are highly variable in both amount and composition. Clearly, if this is so, such losses would have to be

ASTRONAUT C, SUBSTANCE NITROGEN, INPUT SERIES VS. DAY NUMBER
SERIES SYMBOL: CDRIN

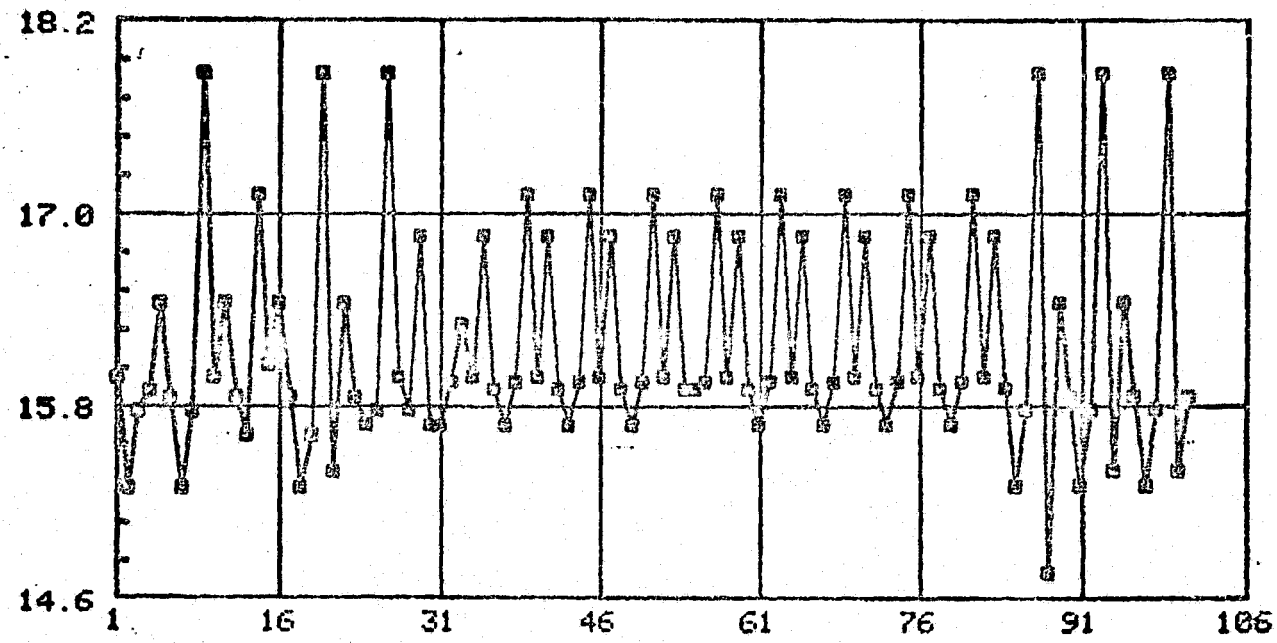



Fig. C1

ASTRONAUT C, SUBSTANCE NITROGEN, URINE SERIES VS. DAY NUMBER
SERIES SYMBOL: CDRUN

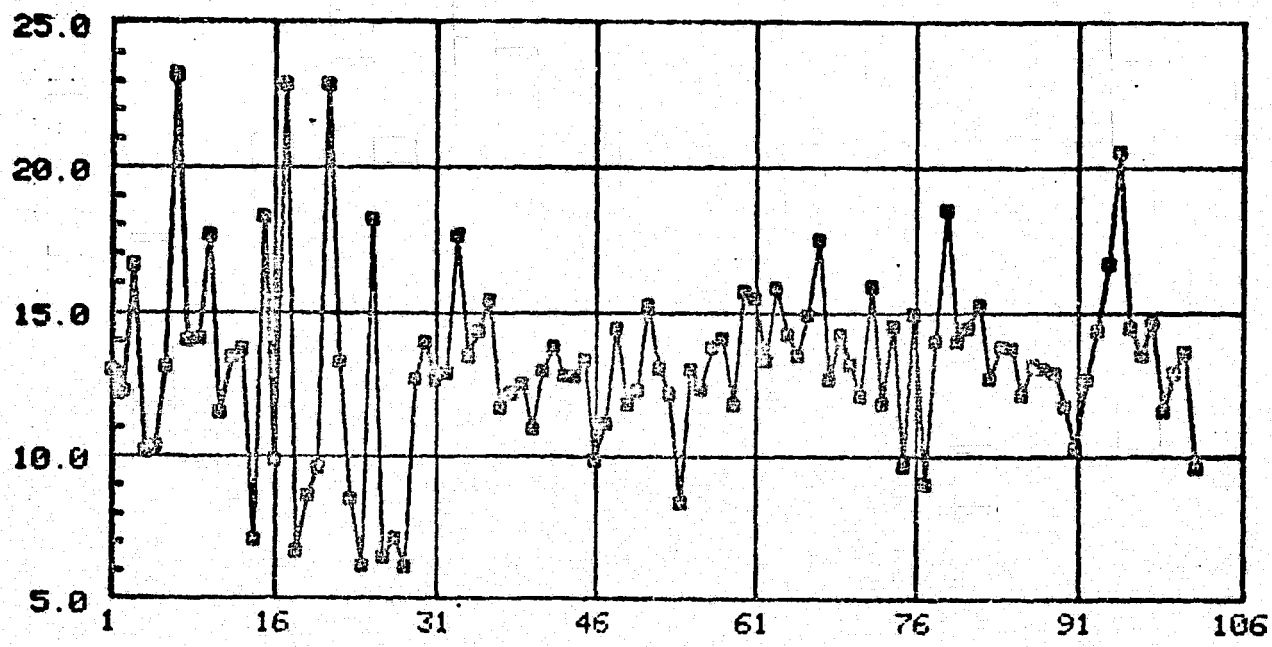



Fig. C2

ASTRONAUT C, SUBSTANCE NITROGEN, FECAL SERIES VS. DAY NUMBER
SERIES SYMBOL: CORFN

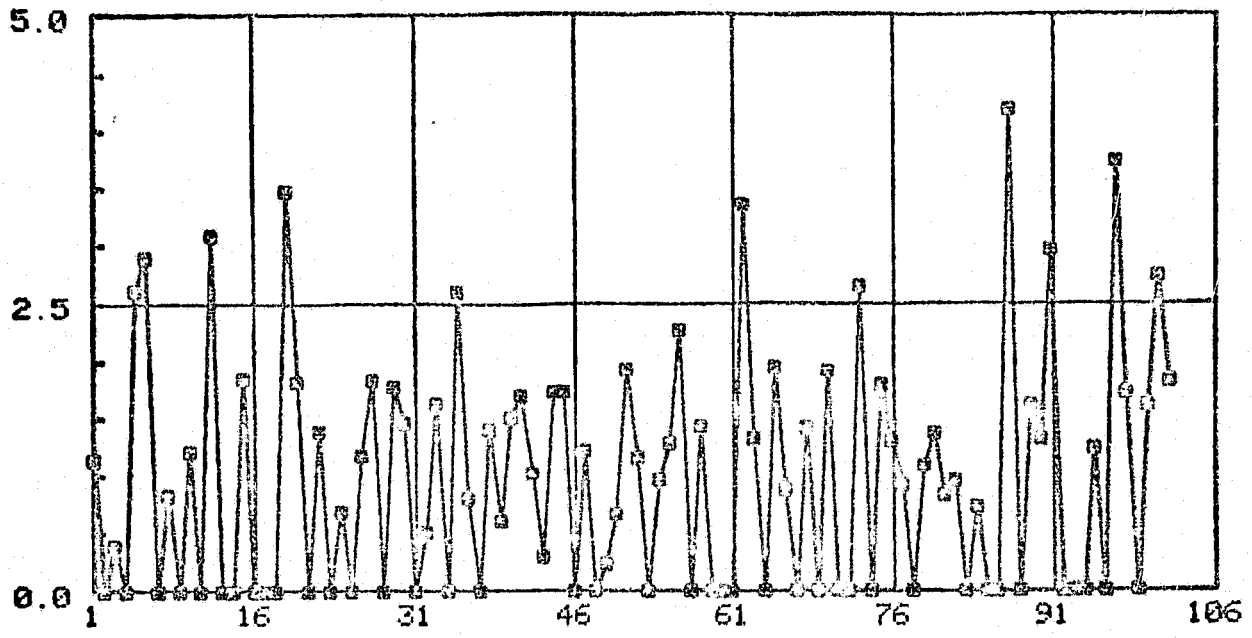



Fig. C3

ASTRONAUT C, SUBSTANCE CALCIUM, INPUT SERIES VS. DAY NUMBER
SERIES SYMBOL: CORICA

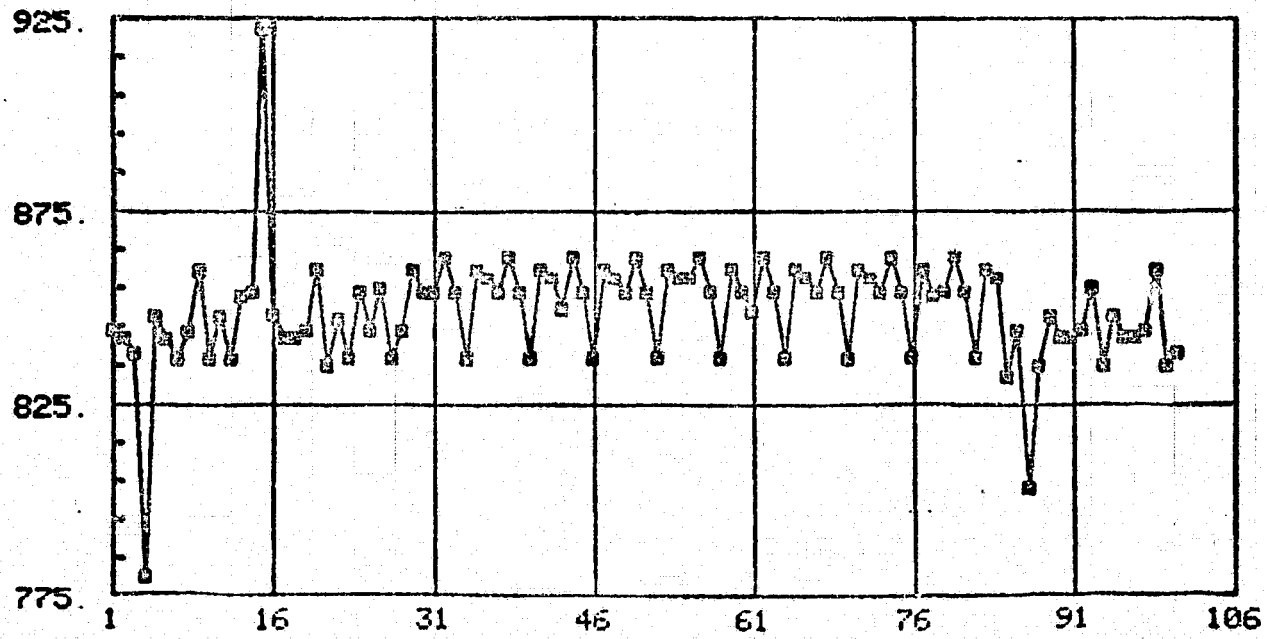



Fig. C4

ASTRONAUT C, SUBSTANCE CALCIUM, URINE SERIES VS. DAY NUMBER

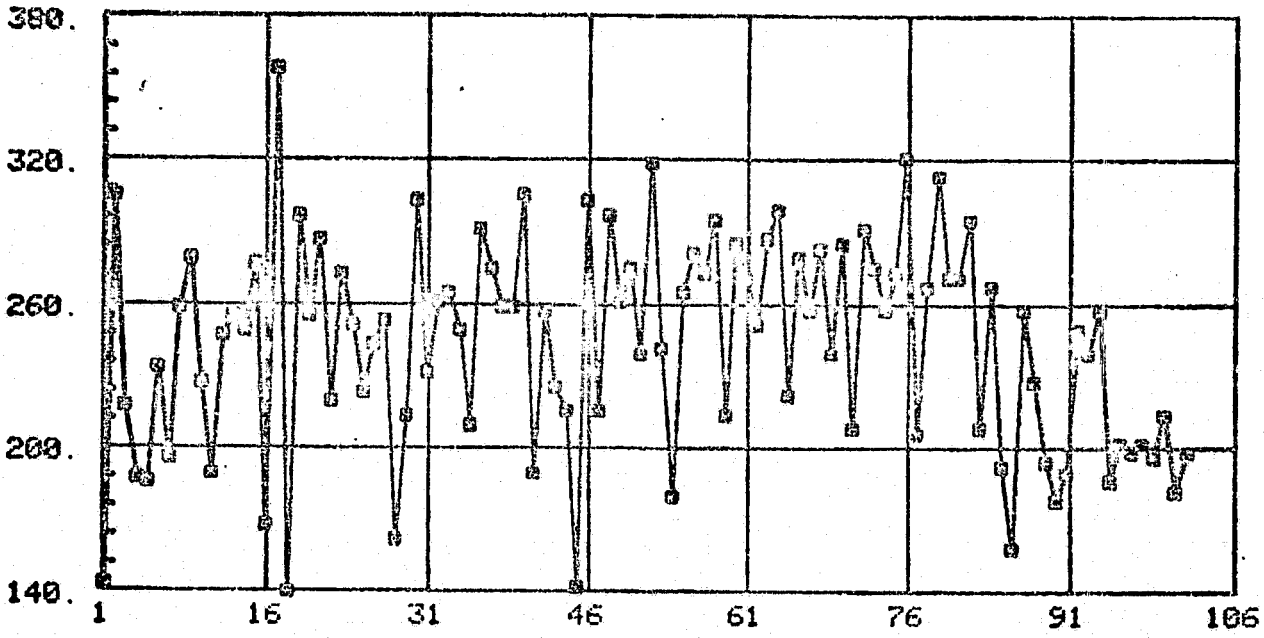



Fig. C5

ASTRONAUT C, SUBSTANCE CALCIUM, FECAL SERIES VS. DAY NUMBER
SERIES SYMBOL: CORFCA

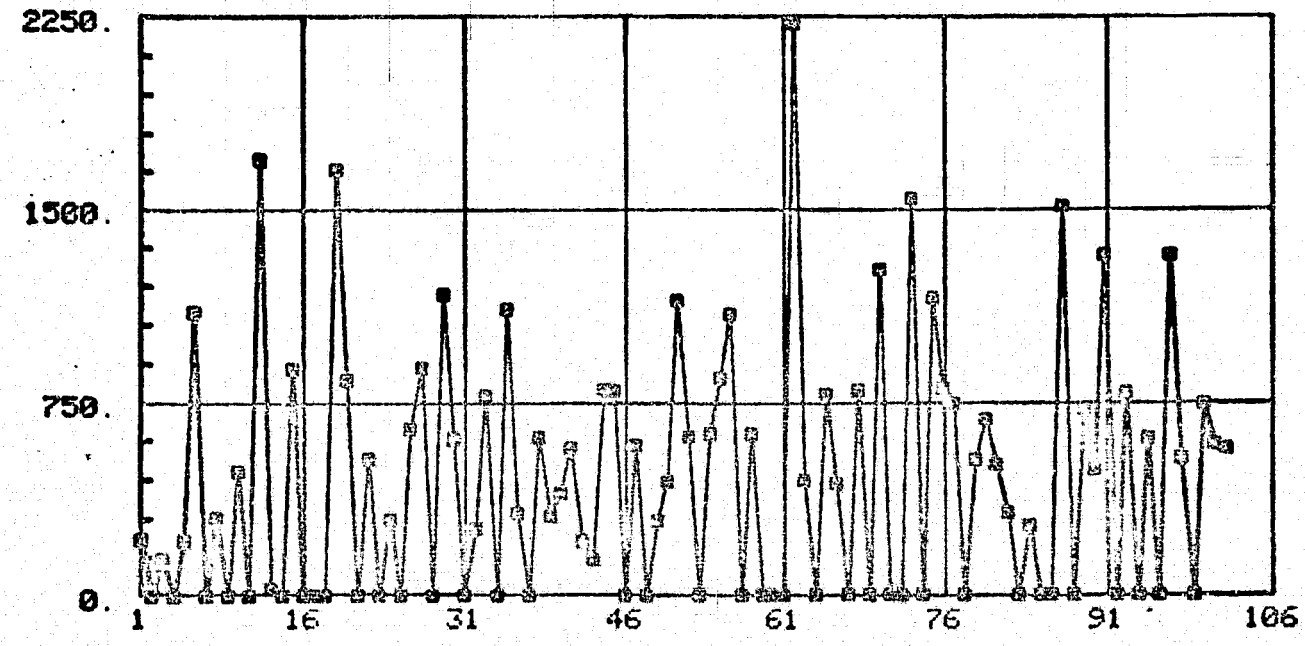



Fig. C6

measured if the balance data are to be meaningful. The application of an "average" correction for dermal excretion, whatever the value chosen, leads to clearly unacceptable results for one or more of the astronauts.

Long-Term Confidence Intervals for NET Retention. The $\hat{E}$(NET) and $\hat{\sigma}^2$(NET) results in Table 1 allow fashioning an approximate confidence interval for NET by

$$(*) \quad \hat{E}(\text{NET}) \pm 2 \left[ \sqrt{\frac{\min(\hat{\sigma}^2(\text{"1-16"}),\ \hat{\sigma}^2(\text{"4-13"}))}{\text{length of NET-series in days}}} \right]$$

where min(a,b) = smaller of a and b. Choosing the smaller variance will give the more optimistic interval; the factor "2" assures that the probability that the interval (*) does not include the "true" NET is small, but the exact probability is not available as not enough information is available to specify the theoretical distribution of NET. Normal probability plots for NET (of which we include only one for SNET N in a later section) suggest that NET is sufficiently gaussian for this approximate factor to yield appropriate intervals but not so for attaching precise tail-probabilities.

These confidence intervals (*) are based on the entire MBE, and as such may perhaps be regarded as "minimum feasible" uncertainties, characterizing their respective MBE's, i.e., intervals that can be achieved through very long MBE's. Table 4 (p.15, Introduction) exhibits these long-term confidence intervals for each of the 9 NET series of Table 1.

For each substance, the three long-term confidence intervals in Table 4 overlap lending credence to the findings. These overlaps, along with the fact that all 9 intervals show $\hat{E}$(NET) > 0 throughout, are consistent with the MBE design bias toward apparent retention.

"Updated" Average Daily Retention. Figs. CUM 1 - CUM 9 exhibit the 9 graphs of the variable CUM of equation (2) for the three astronauts and the three nutrients under study as against the day number of the MBE's. Gaps in some nitrogen data allow us to continue two of the CUM nitrogen series only to day 74; the other 7 extend to day 101 (for "C" and "S") or 102 (for "P").

Figs. CUM 10 - CUM 12 plot the CUM series for all three astronauts on the same axis, for each of the three substances: these plots exhibit the gradually increasing consistency in the average NET's for each astronaut as their MBE's progressed.

The upturn of SCUMK on days 89-101 is spurious being due to a sudden increase in SPTICa during these final days to which SPT evidently had not the time to acclimatize.

The CUM plots have the general appearance of oscillating wildly for the first 3 or 4 diet periods (18-24 days) and then assuming a smooth character that gradually either increases or decreases, often with a perceptible nonlinearity (concavity or convexity). A qualitative look as to how far along into the MBE's the CUM "stabilizes" suggests that comparatively little new information is added by, say, the second half of the MBE; the implication regarding optimal MBE design will be discussed in a following section on conjunction with the issue of distinguishing true, or biologically-based, trend. We wish now to discuss the interpretation of the concavity noted above as an artificial trend.

Artificial Trend. The concavity, giving an appearance of trend in average NET over time, is at least partly an artifact of the averaging process. This effect can be understood as being due to a term in $CUM_k$ equal to $m\Delta/k$ where $\Delta$ is the difference between the average of the daily NET's up to day m

and this average after day m, where m is small compared to k.

Fig. CUM 13 emphasizes this point. This plot shows a not atypical realization of $C_k$ versus k, where

$$C_k = \sum_{i=1}^{k} N_i$$

and $N_i$ is a random sample of size 100 from a gaussain distribution with mean 0 and variance 1.

We do not discuss how departures from gaussianity of the NET distributions are reflected in the analogous $CUM_k$ versus k graphs; for the present purpose we find the empirical distributions of the NET variables to be sufficiently gaussian as shown by normal probability plots. Fig. NORMN is such a plot for SNET N (the SN MBE was a comparatively variable MBE). Note: Although the plot is close enough to a straight line for the present purpose, it has perceptibly heavy tails, a typical finding for the other NET plots (which we do not include) and this accounts for our qualifications about the exact confidence levels associated with the intervals computed below.

Coefficients of Variation per MBE. If we regard for the moment the estimates of NET variance from periods 4-13 in Table 1 as estimates available to the MBE's prior to the MBE, i.e., as known characteristic numbers for this experimental procedure (and also for this subject in the present use), then we may obtain estimates of the accuracy of the nine MBE's corresponding to each subject and each nutrient as a function of MBE length in days as follows. The coefficient of variation for each MBE on day d is

$$CV_d = \frac{\sqrt{\hat{\sigma}^2/d}}{\hat{E}_d(NET)}$$

where $\hat{\sigma}^2$ is the variance estimate as above from Table 1 and

$$E_d(\text{NET}) = \text{CUM}_d = \frac{\sum_{i=1}^{d} \text{NET}_i}{\alpha}$$

(Compare (*) where d = maximum day number). Figs. CV 1 - CV 9 show $CV_d$ versus d, i.e., the uncertainty in the MBE as a function of experiment's duration in days, using the "updated mean" $\hat{E}_d$ and the "prior variance" $\hat{\sigma}^2$. (An "updated variance" might involve successive analyses of variance for successive periods, but we do not view this as an important ramification for the present study.) Because $\hat{E}_d$ (SNET N) $\to 0$ as $d \to 101$, the CV fluctuates wildly; for this case we display in Fig. CV1a the more stable statistic $CV^1$ where

$$CV_d^1 = \frac{\sqrt{\hat{\sigma}^2}}{\hat{E}_{101}(\text{SNET N})}$$

which regards the limiting NET as known in advance of the MBE (as CV regards the variance estimate only). For this case $CV^1$ is proportional to $d^{-\frac{1}{2}}$ exactly which accounts for its smoothness. Fig. CV3a exhibits CV for PN on an exceptional scale; as PCVN does not take on reasonable values except for very large day numbers, we includ $PCV^1N$ as well. (Note: Because CV (and $CV^1$) vary, approximately and exactly respectively, like $d^{-\frac{1}{2}}$, where d is the day number, the asymptotes of the CV graphs are all, of course, the CV = .00 - line.) The CV's for K are very consistent and, in view of the graphs' stabilities suggest that together they constitute a fairly accurate determination. This, however, is only to say that the data are rather consistent. We have previously noted that the absolute retentions of potassium are so high as to be meaningless and large unmeasured losses must occur. The retentions for calcium do not yield as consistent a picture and those for nitrogen are least adequately characterized. These "updated"

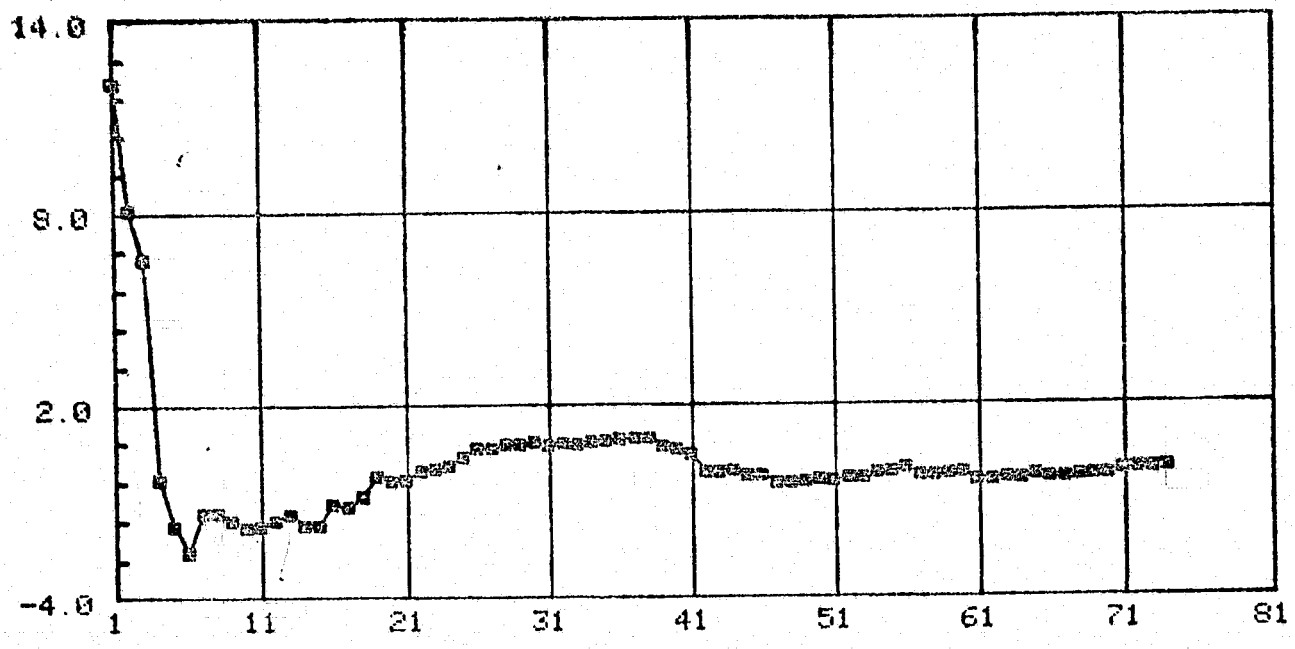



Fig. CUM 1. Plot of cumulative nitrogen retention for astronaut "S".

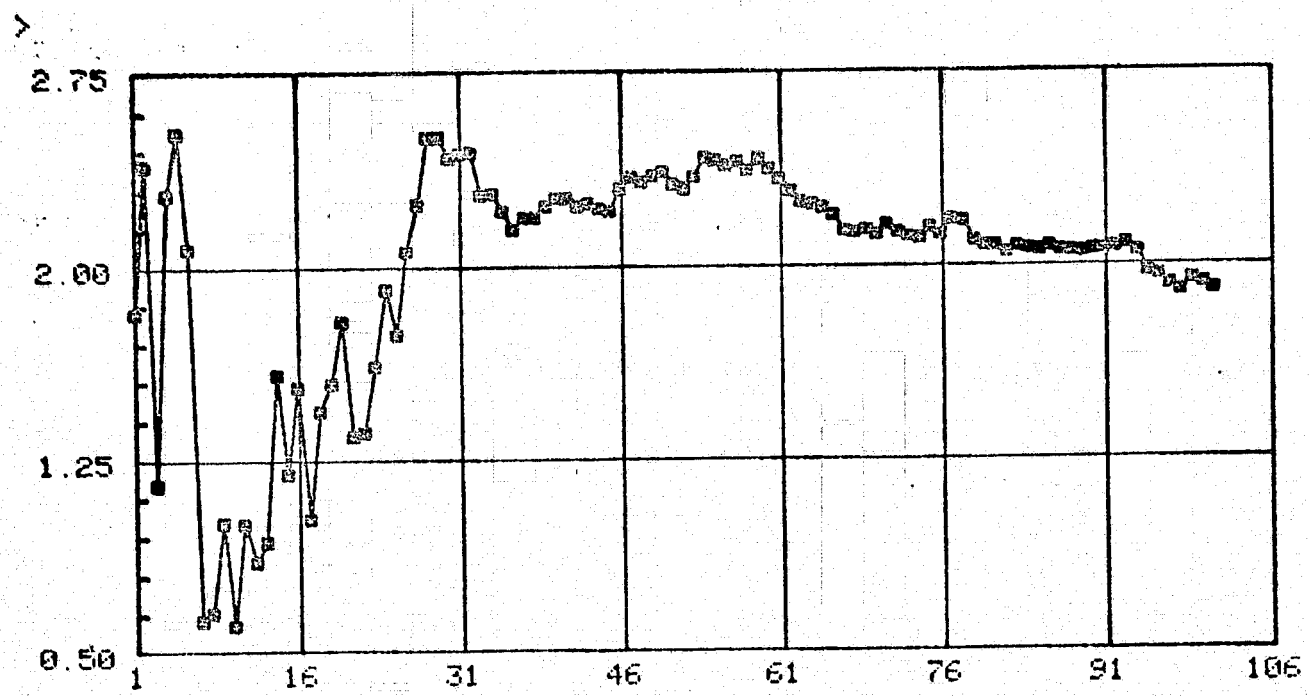



Fig. CUM 2. Plot of cumulative nitrogen retention for astronaut "C".

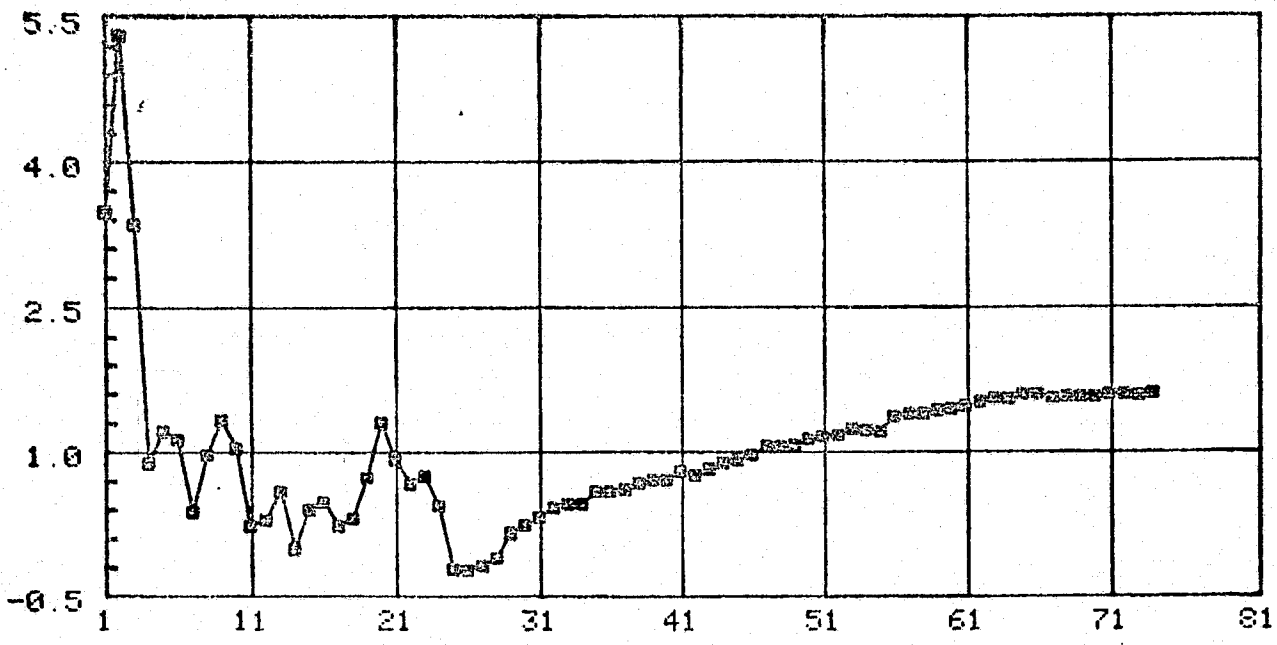



Fig. CUM 3. Plot of cumulative nitrogen retention for astronaut "P".

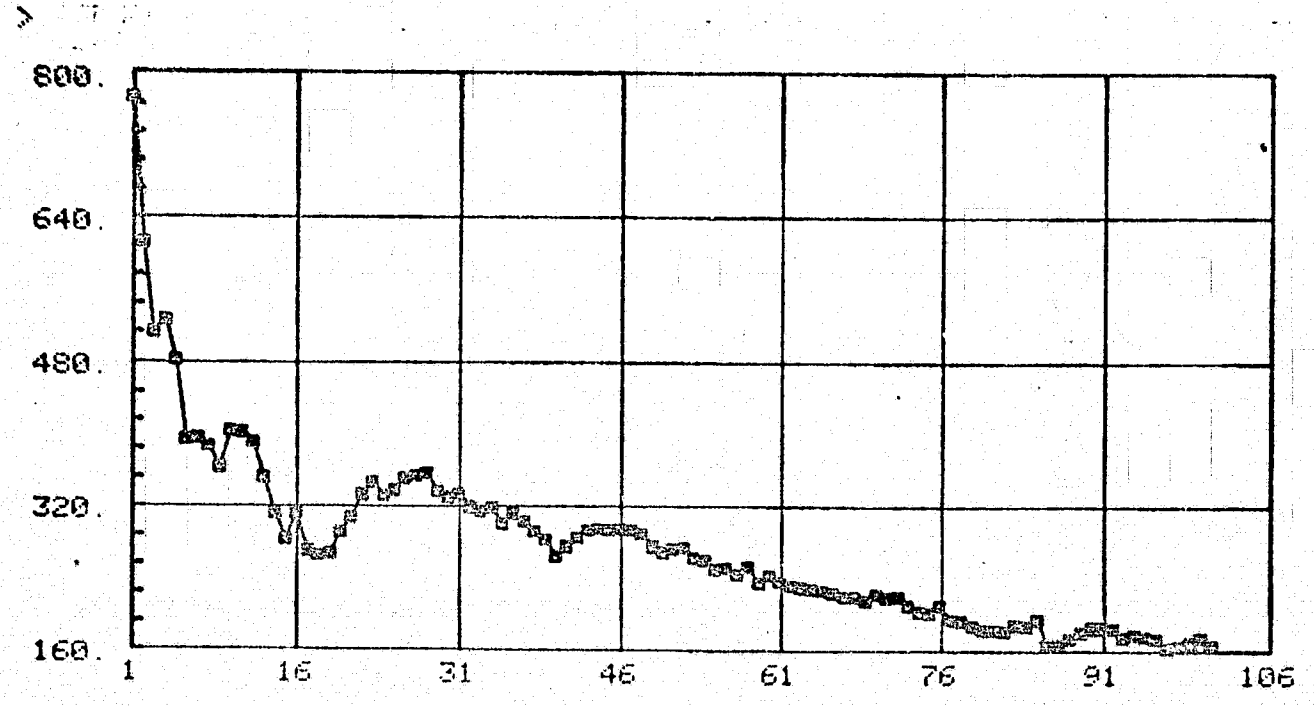



Fig. CUM 4. Plot of cumulative calcium retention for astronaut "S".

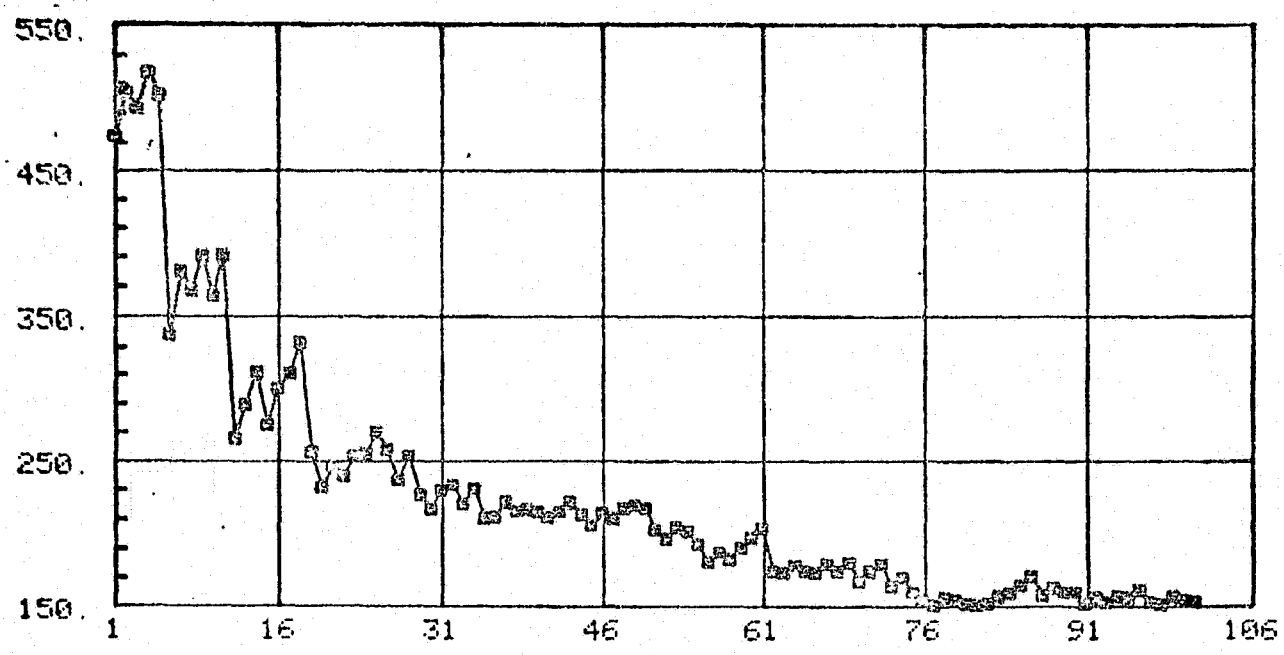



Fig. CUM 5. Plot of cumulative calcium retention for astronaut "C".

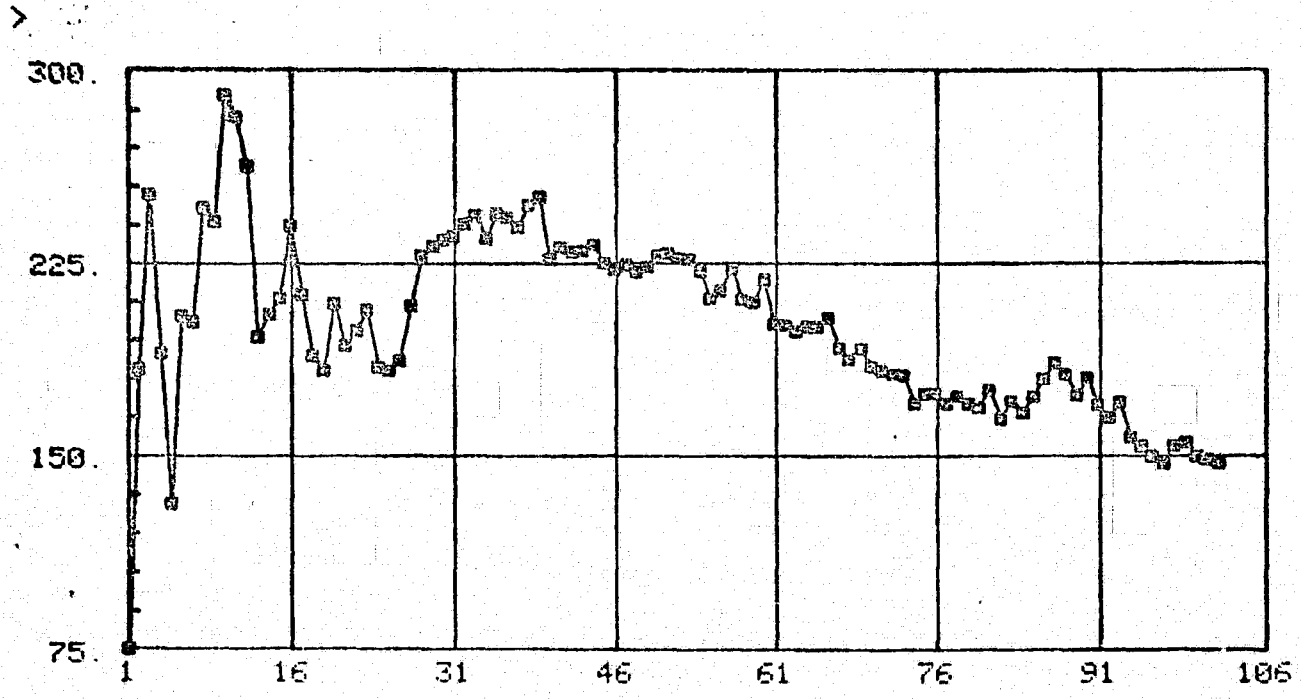



Fig. CUM 6. Plot of cumulative calcium retention for astronaut "P".

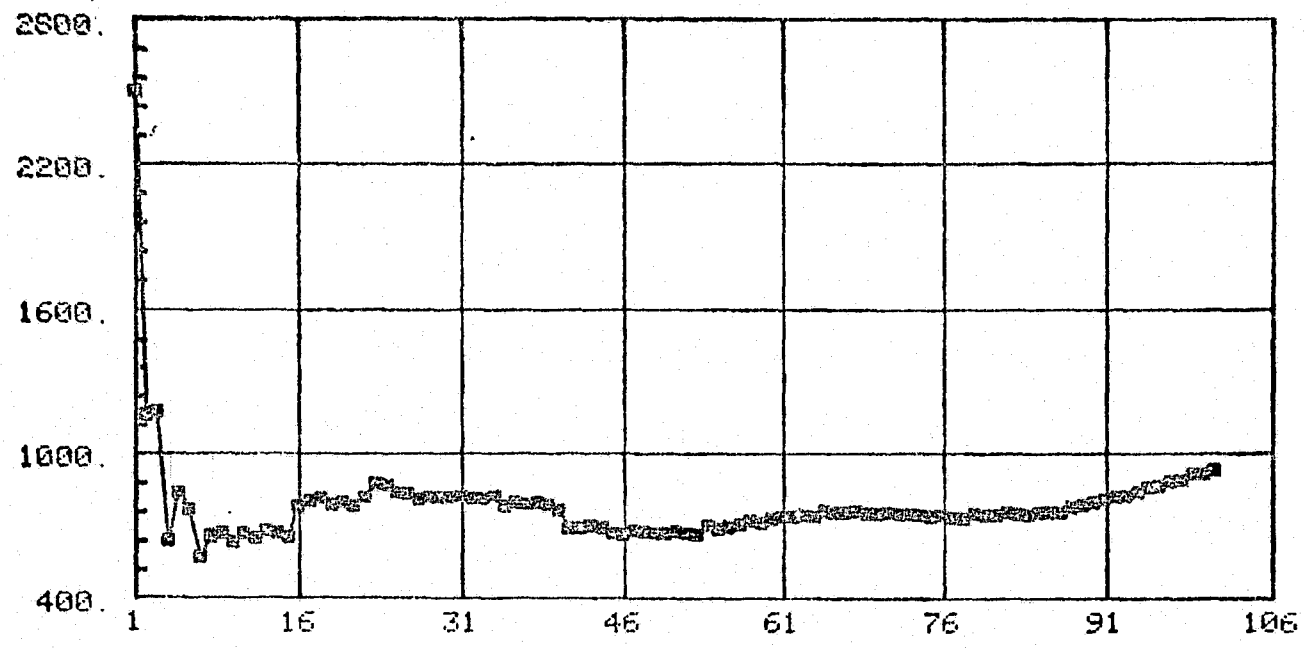



Fig. CUM 7. Plot of cumulative potassium retention for astronaut "S".

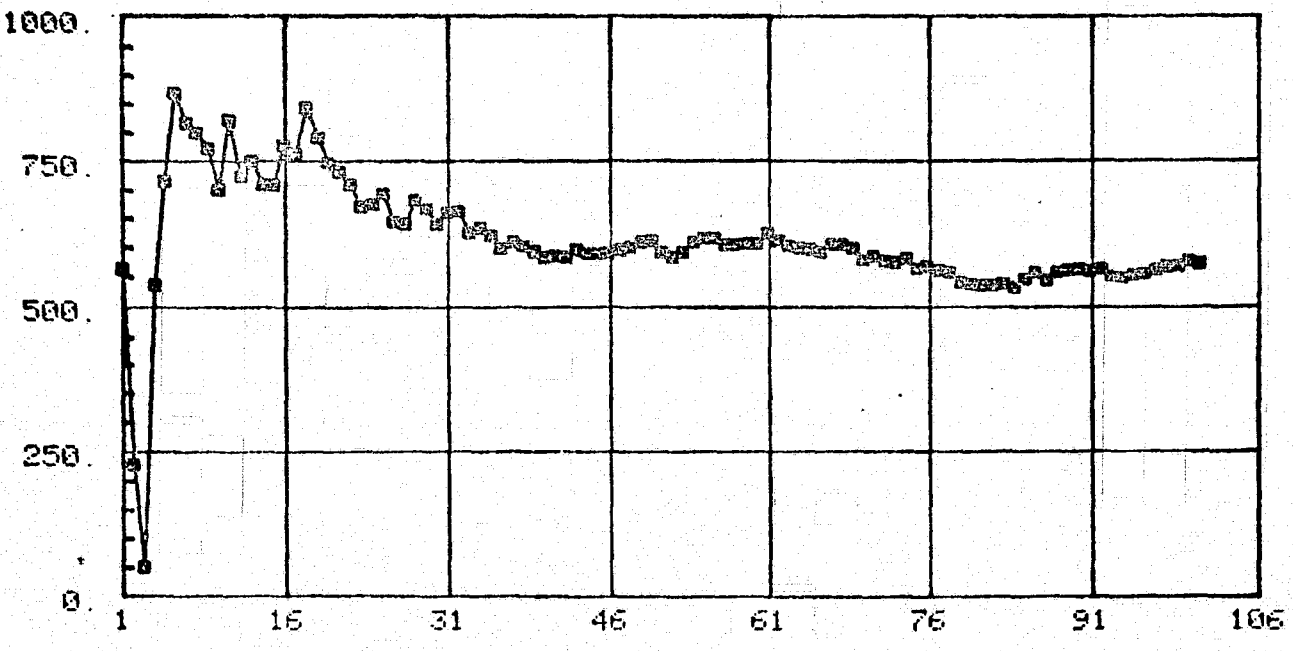



Fig. CUM 8. Plot of cumulative potassium retention for astronaut "C".

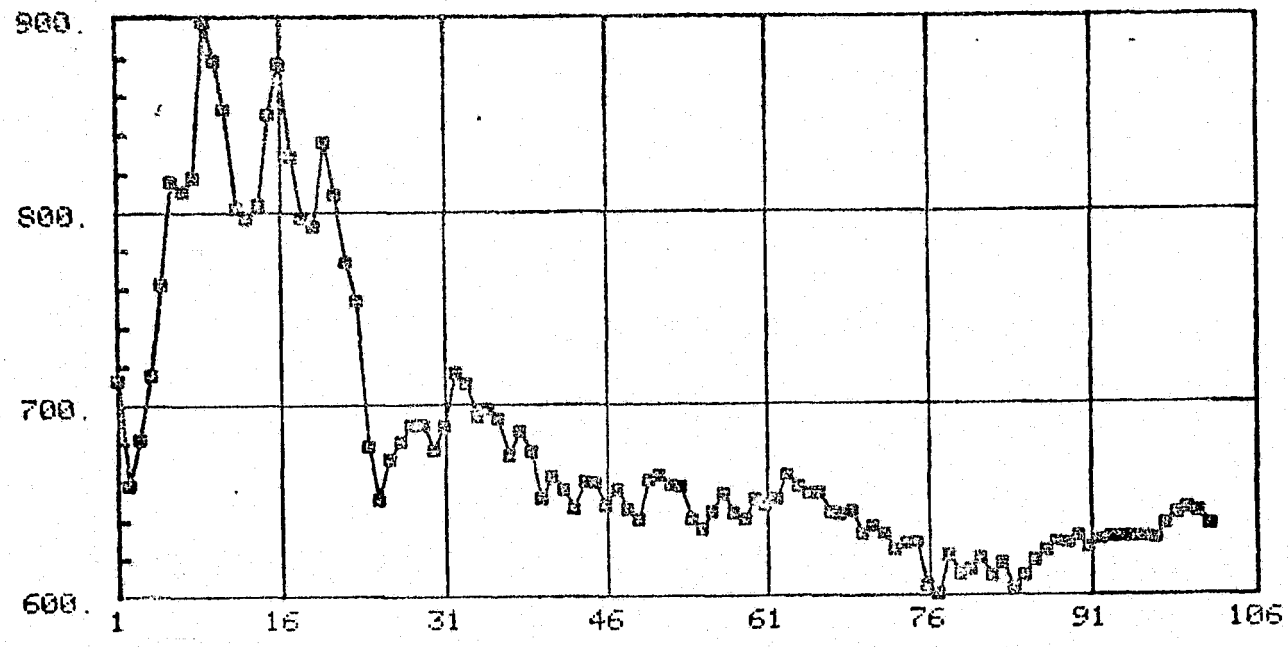



Fig. CUM 9. Plot of cumulative potassium retention for astronaut "P"

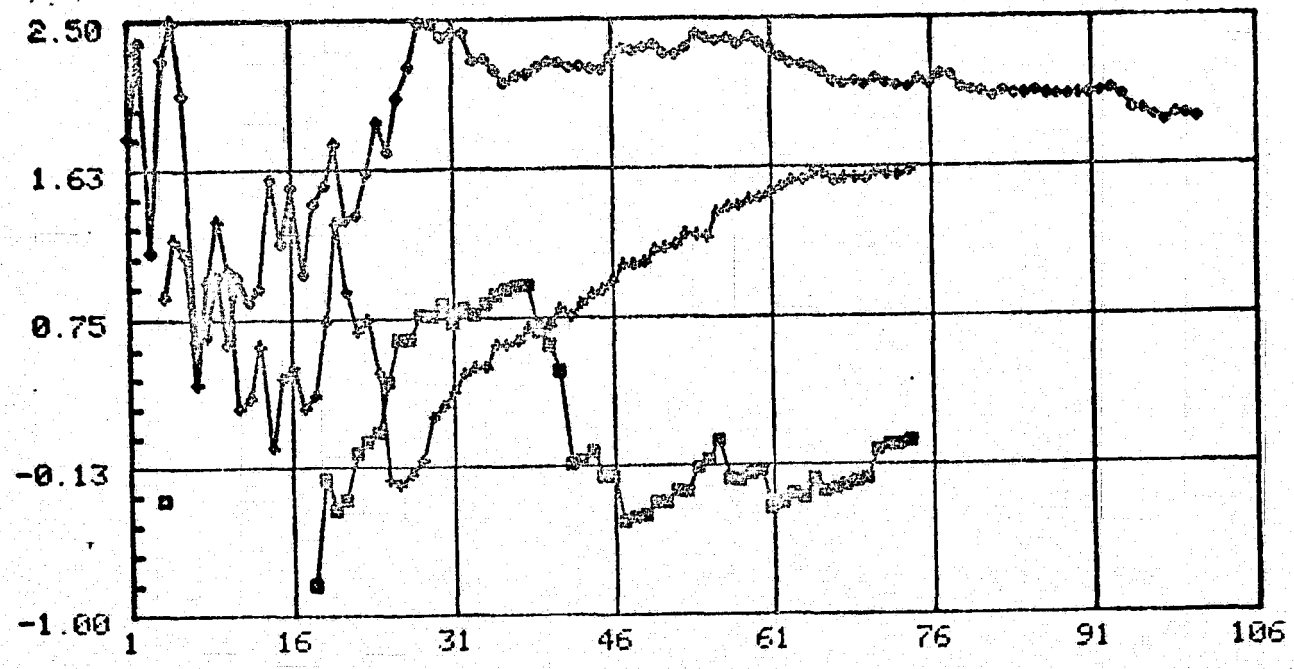



Fig. CUM 10. Cumulative nitroten retention for the 3 astronauts over time.

■ #1 SCUMN
● #1 CCUMN

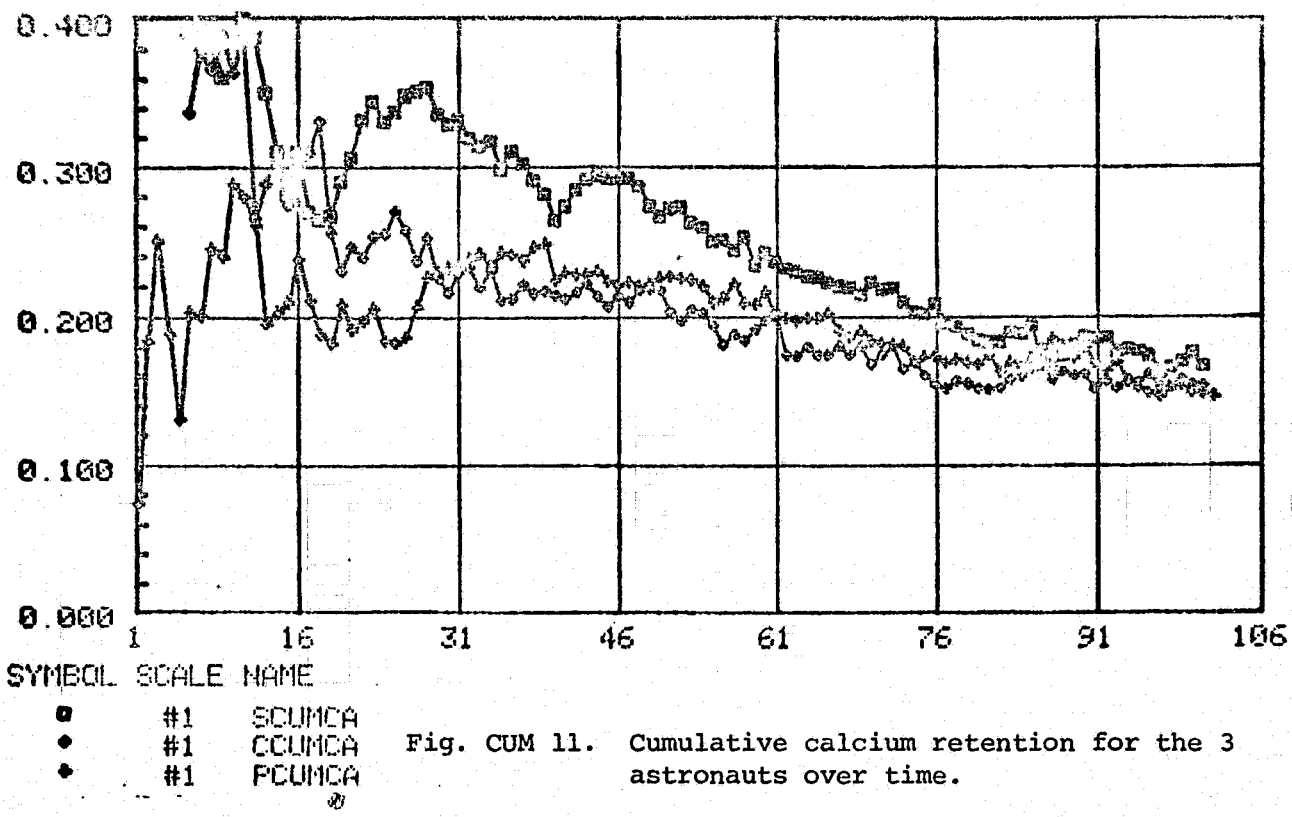



Fig. CUM 11. Cumulative calcium retention for the 3 astronauts over time.

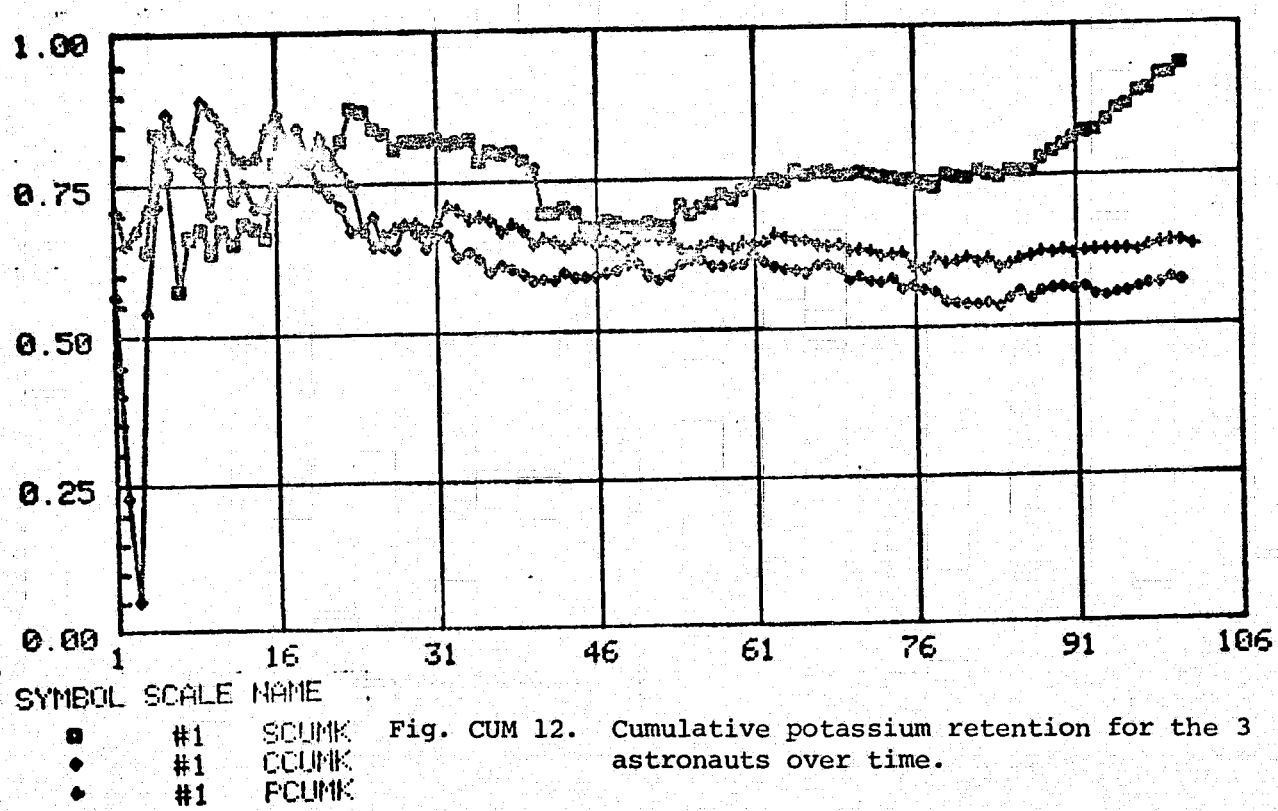



Fig. CUM 12. Cumulative potassium retention for the 3 astronauts over time.

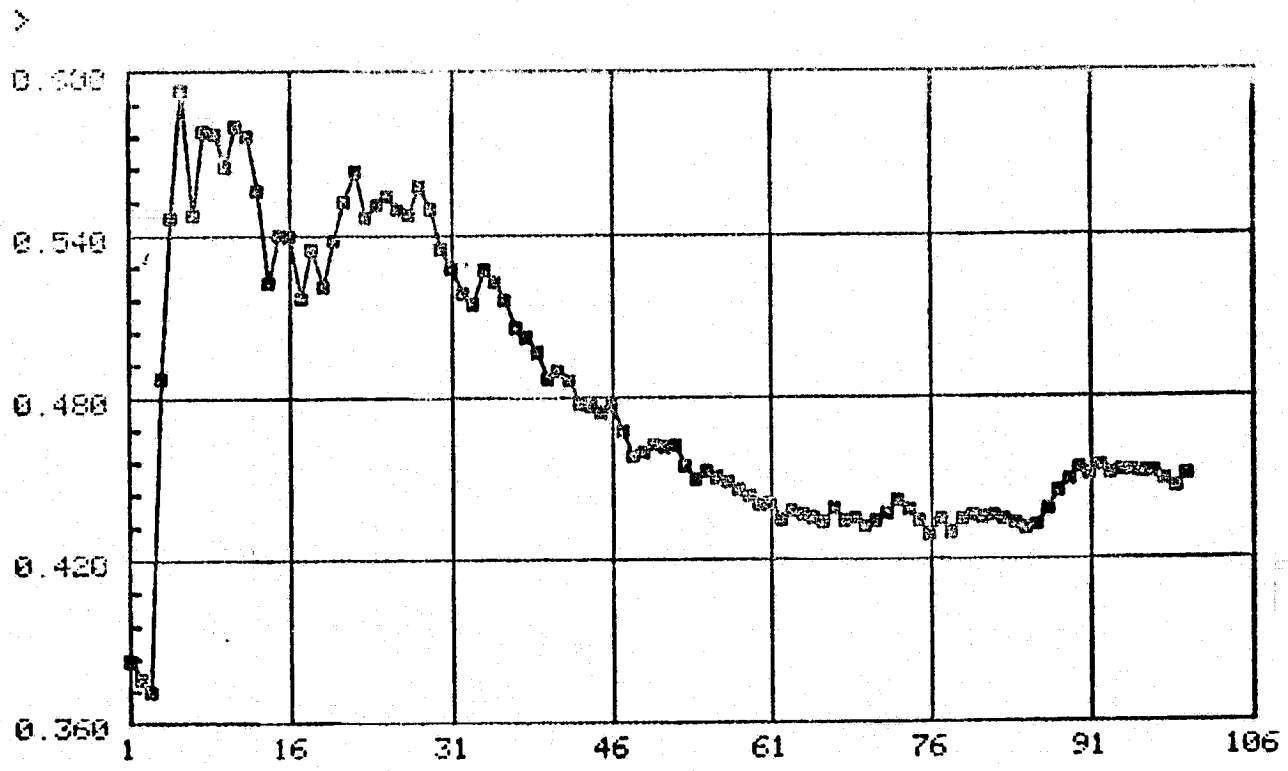



Fig. CUM 13. Theoretical plot of cumulative retention (see text, p.28).

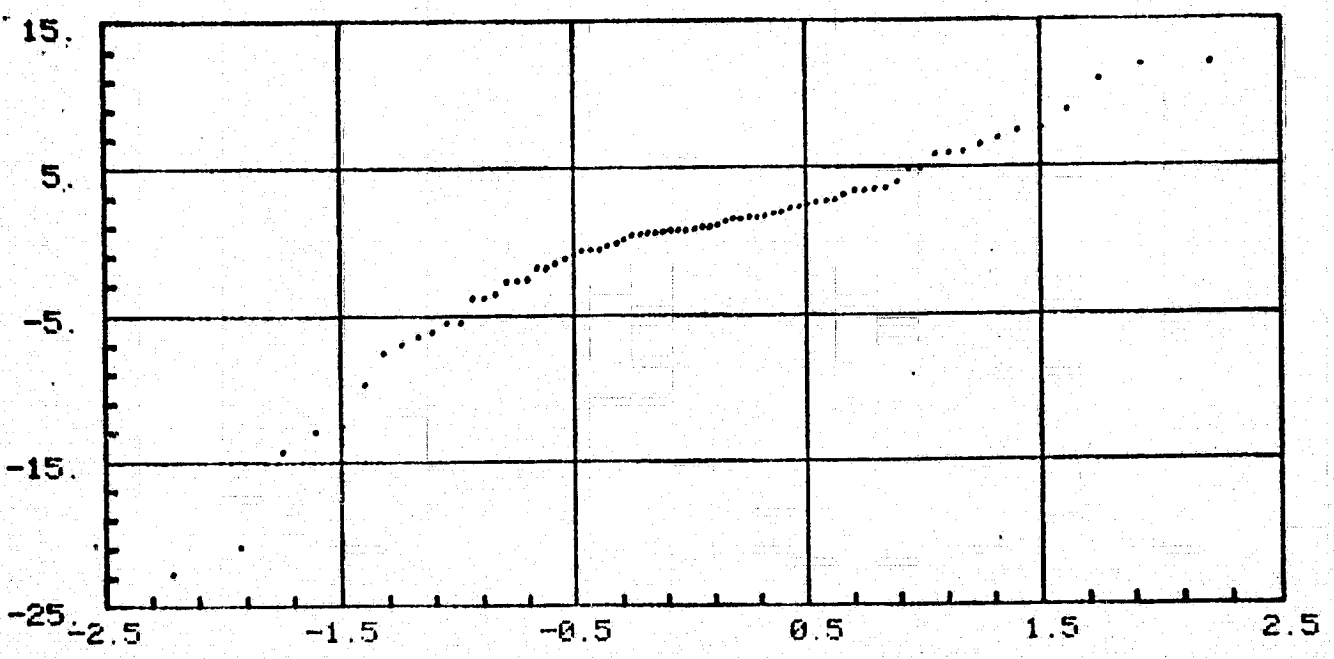



Fig. NORM N. Normal probability plot of NET nitrogen in astronaut "S".

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
US MBE DURATION IN DAYS, ASTRONAUT S, MINERAL N

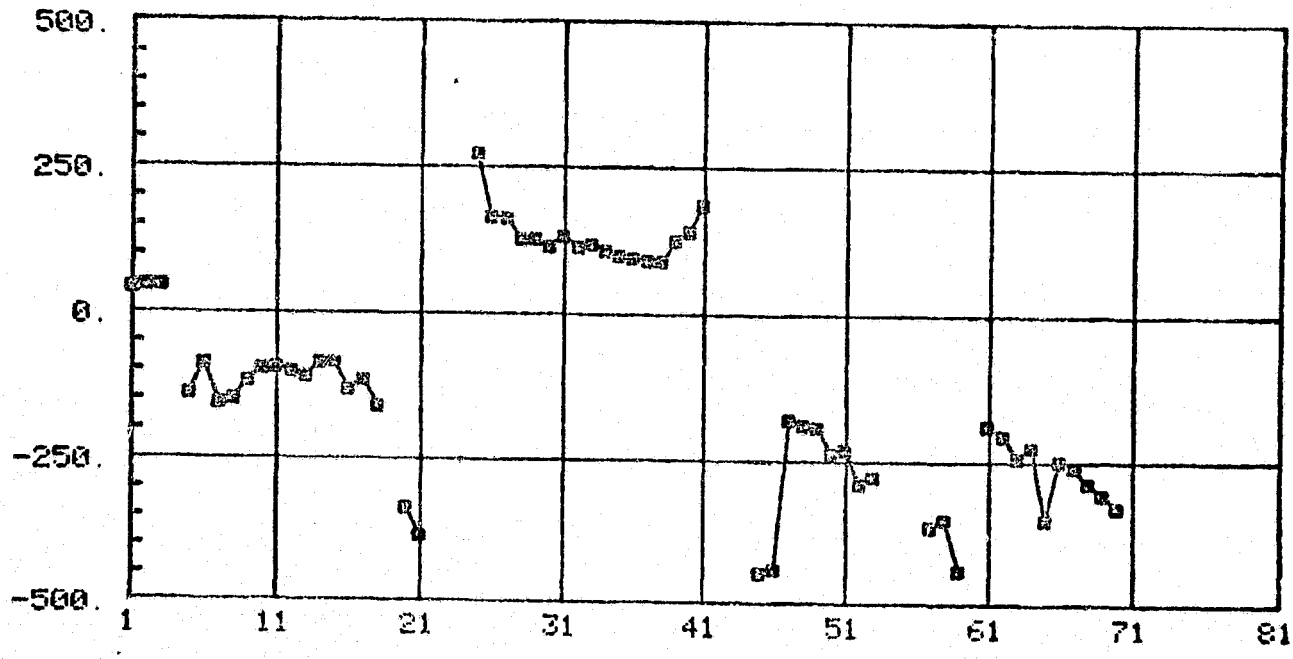



Fig. CV 1

PERCENT COEFFICIENT OF VARIATION VS. MBE DURATION IN DAYS,
ASTRONAUT S, MINERAL N

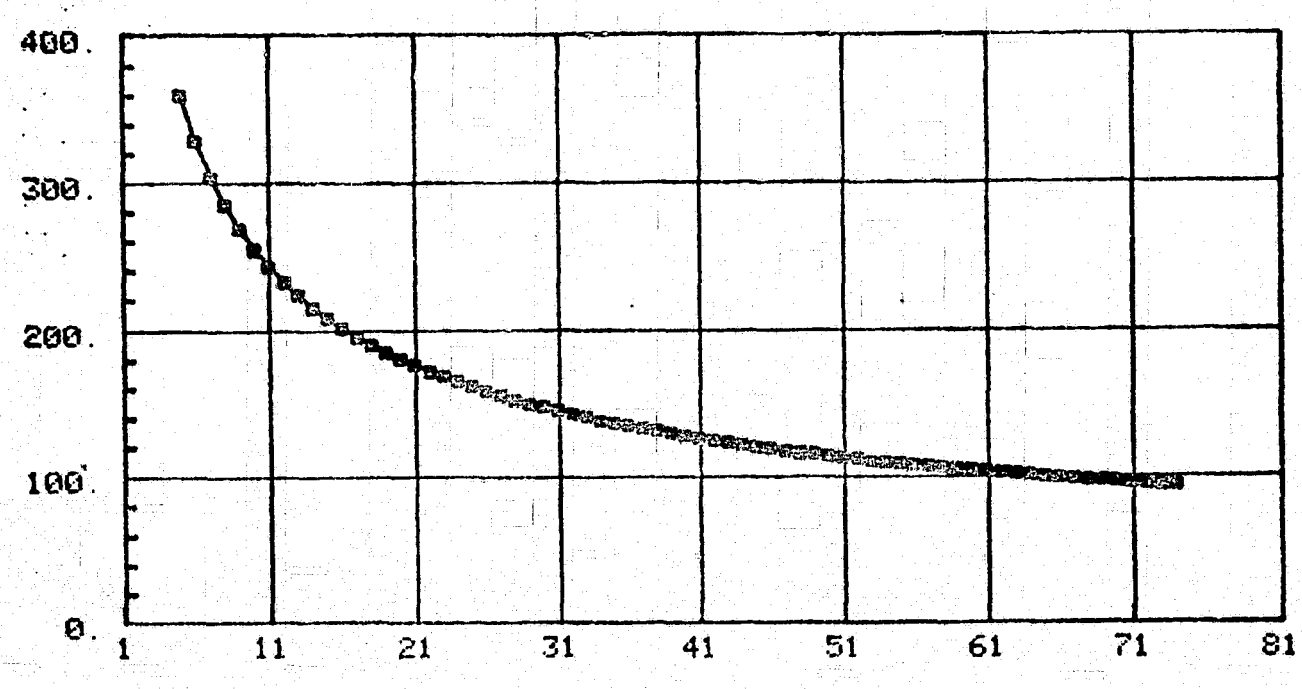



Fig. CV 1a

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT C, MINERAL N

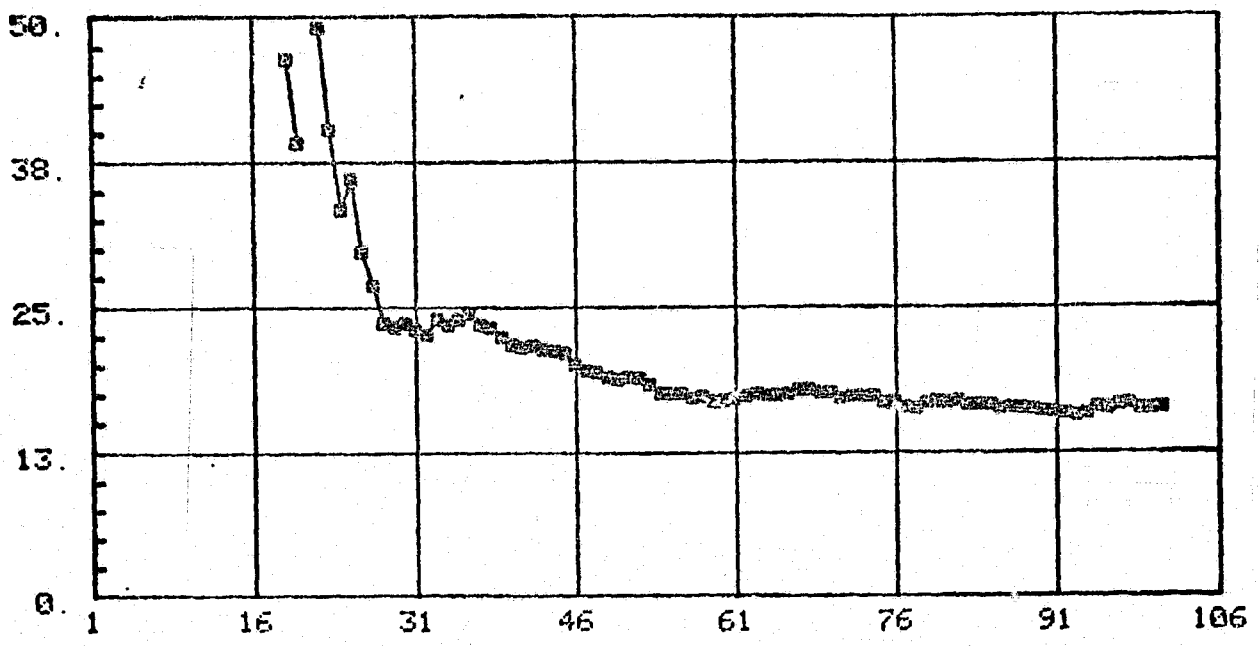



Fig. CV 2

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT P, MINERAL N

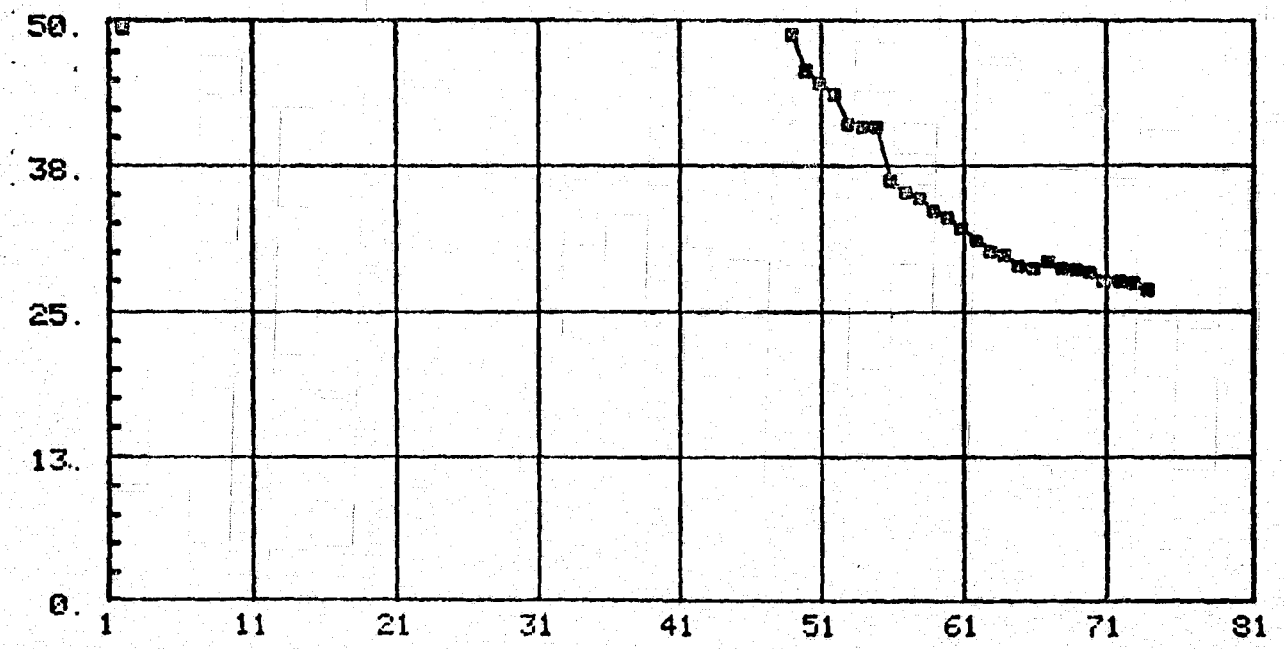



Fig. CV 3

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT P, MINERAL N

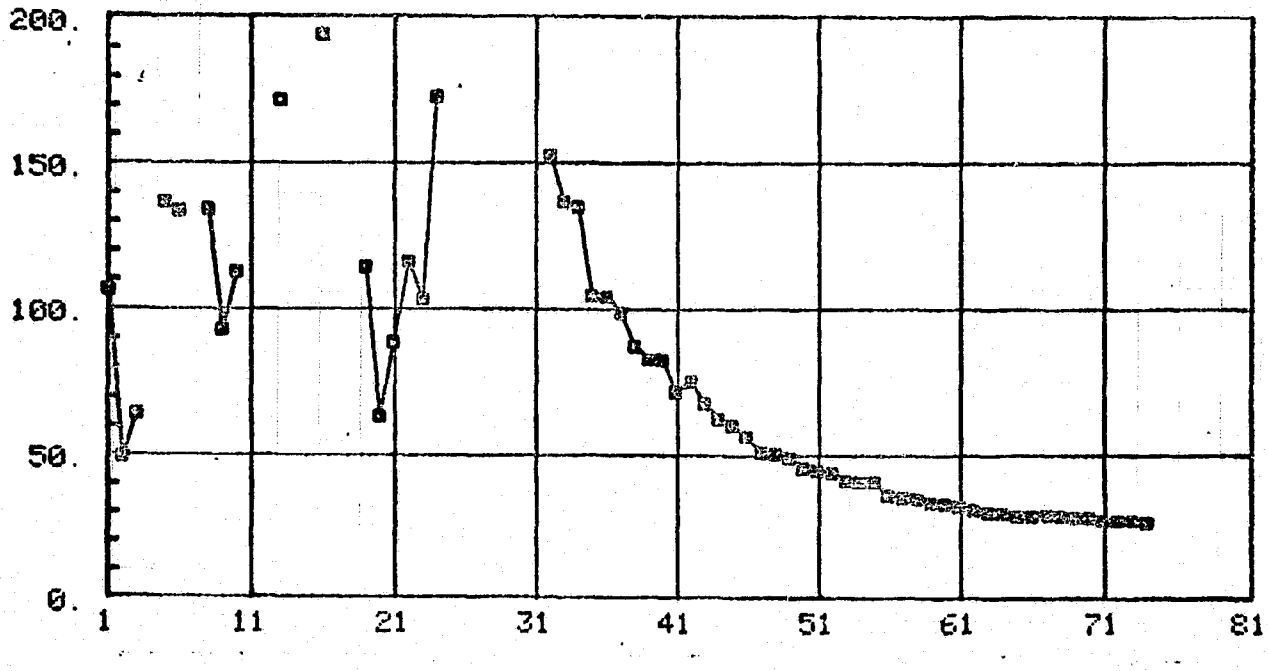



Fig. CV 3a

PERCENT COEFFICIENT OF VARIATION VS. MBE DURATION IN DAYS,
ASTRONAUT P, MINERAL N

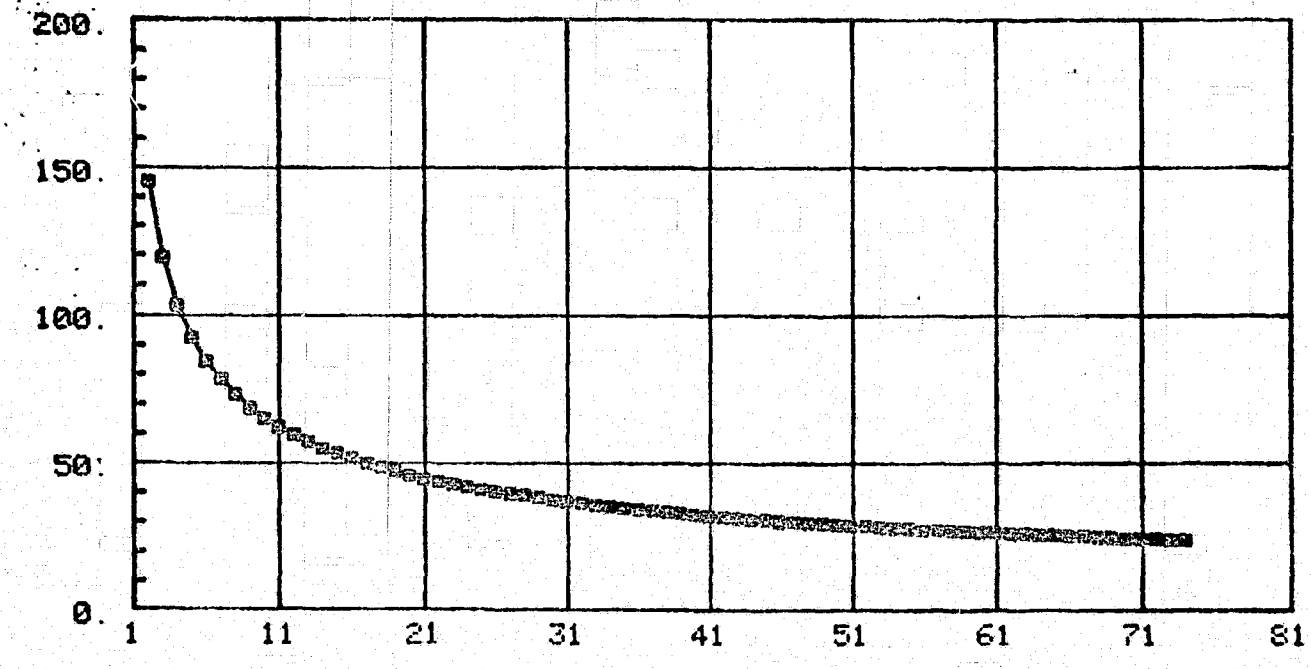



Fig. CV 3b

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT S, MINERAL CA

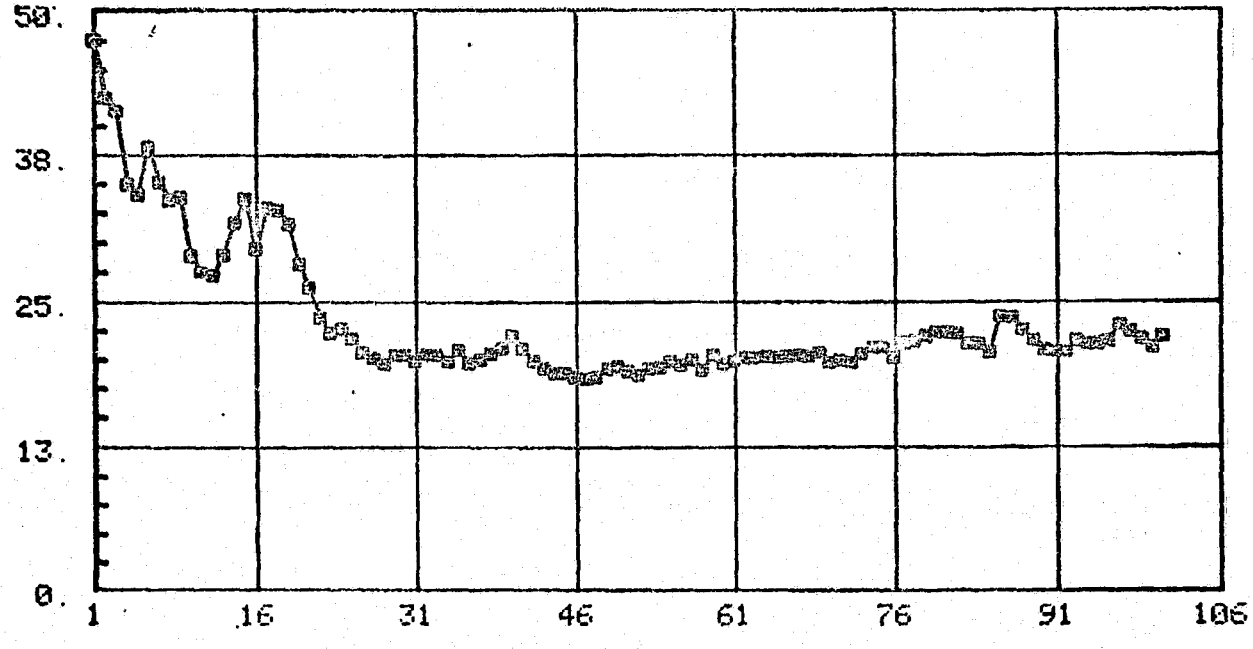



Fig. CV 4

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT C, MINERAL CA

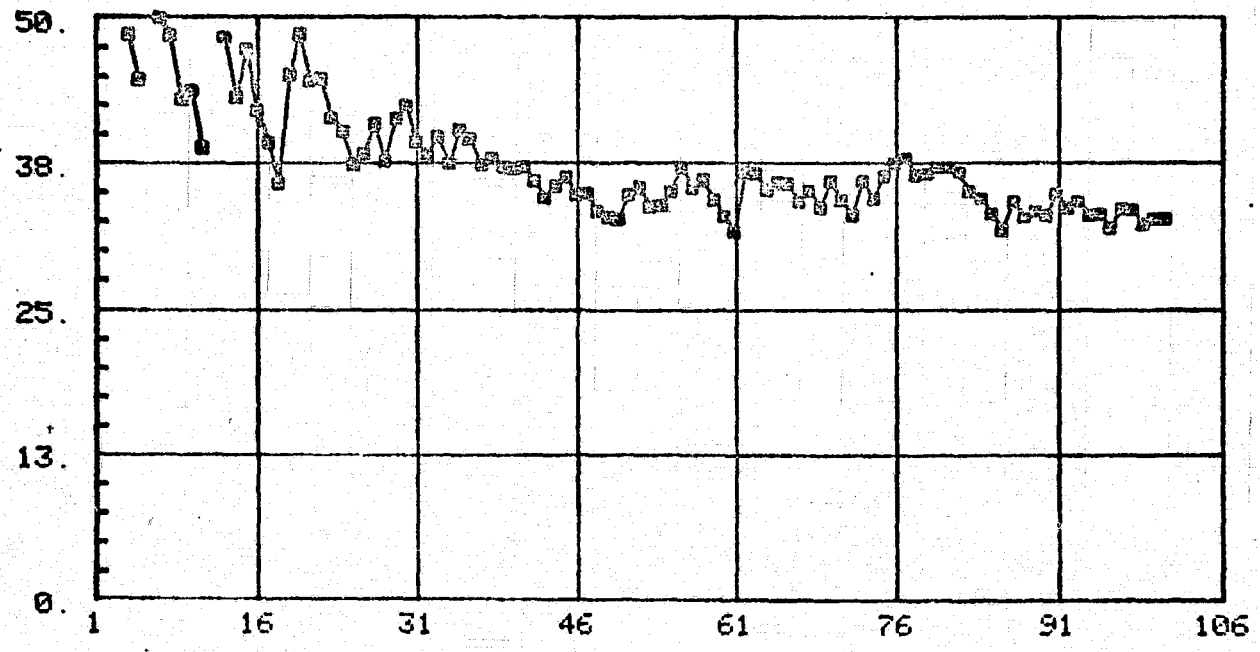



Fig. CV 5

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT P, MINERAL CA

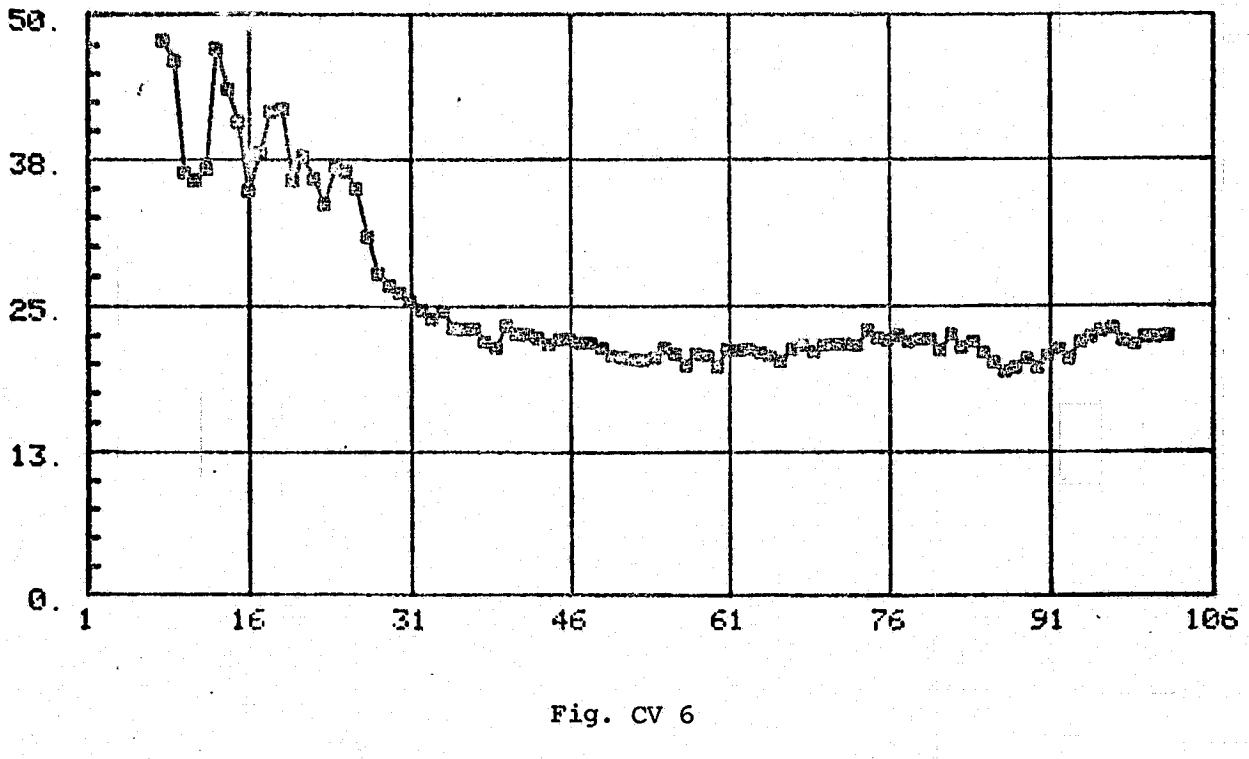



Fig. CV 6

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT S, MINERAL K

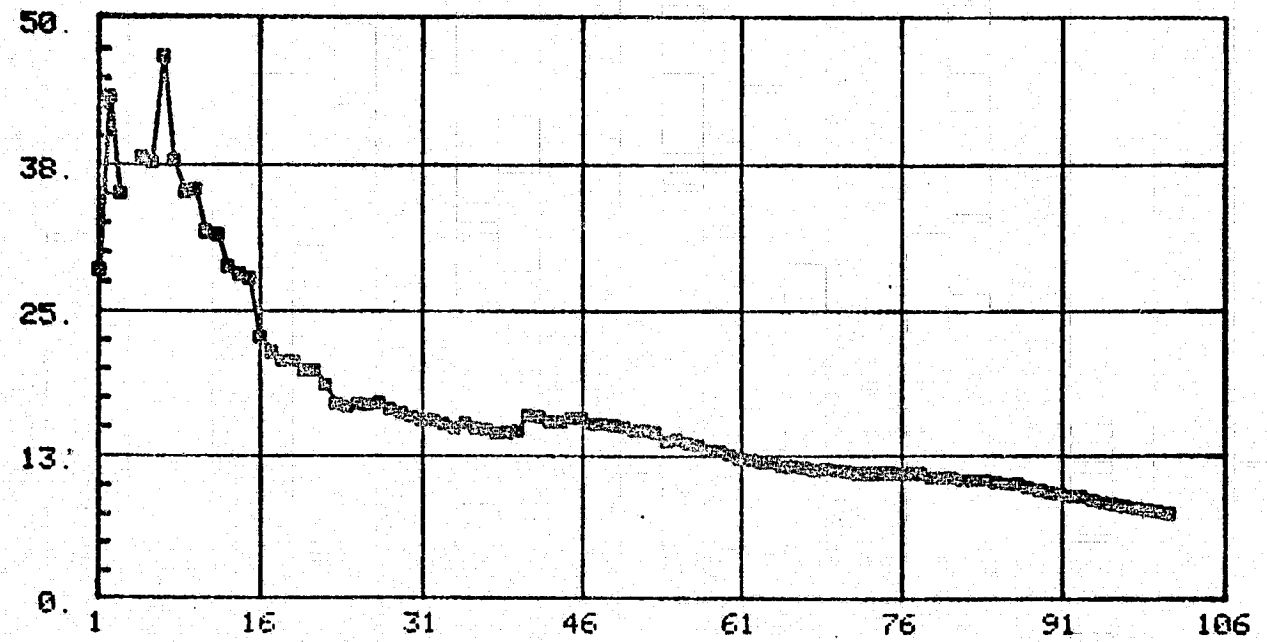



Fig. CV 7

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT C, MINERAL K

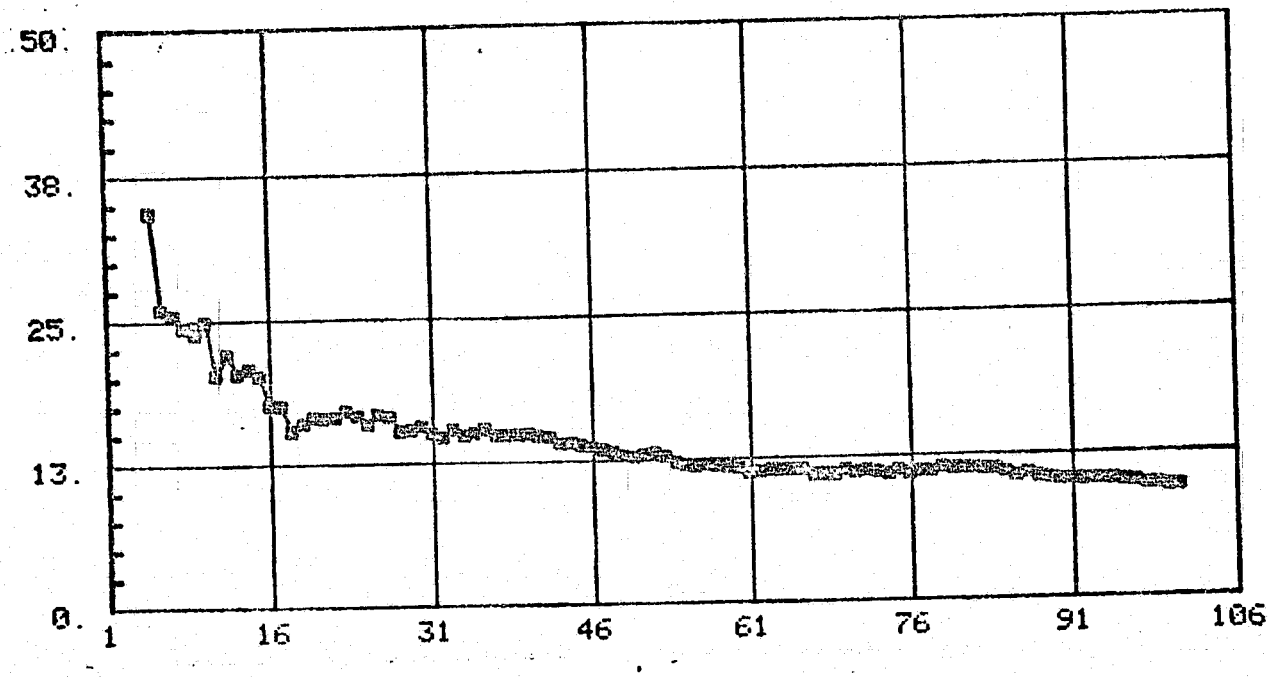



Fig. CV 8

RUNNING COEFFICIENTS OF VARIATION OF NET=I-U-F, IN PERCENT,
VS. MBE DURATION IN DAYS: ASTRONAUT P, MINERAL K

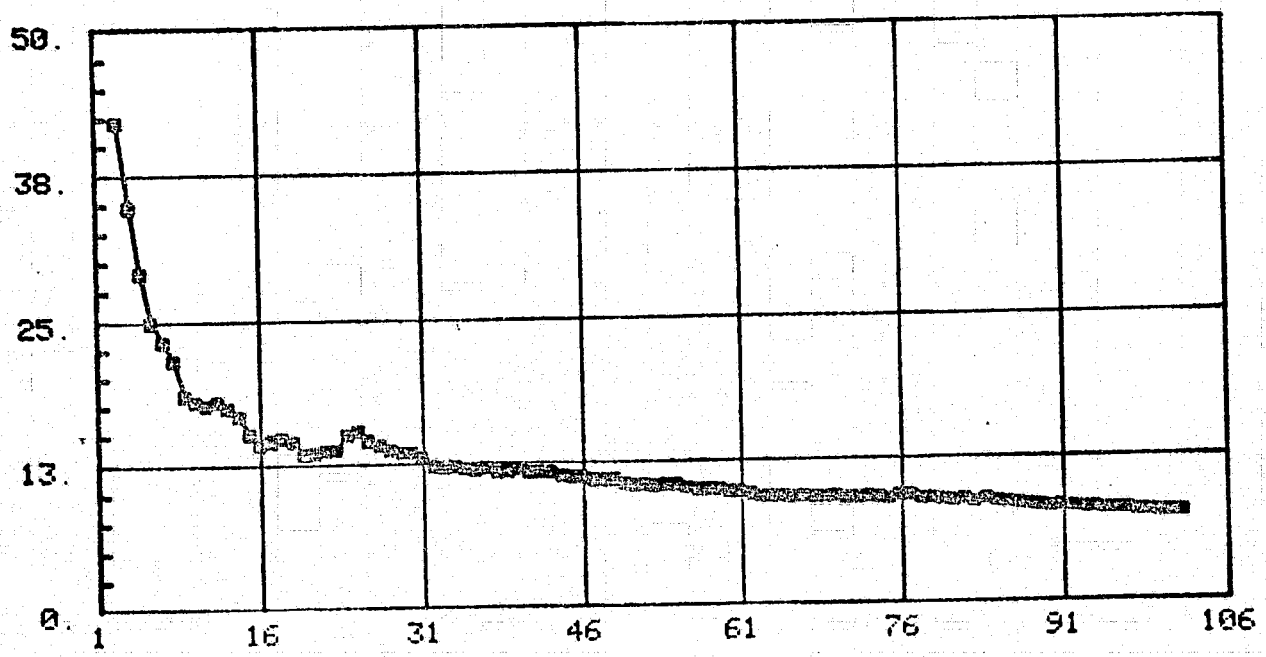



Fig. CV 9

CV characteristics accord with those of the limiting confidence intervals in Table 4.

The "Updated" Confidence Interval for Net Retention. The "updated confidence intervals" (equation 7) for average net retention, E(NET), as a function of MBE duration, may be obtained in view of (3) and (4) from the $CUM_k$ and the value of $\hat{\sigma}^2(NET)$ in Table 1, for any astronaut and any substance, as

$$\frac{\sum_{i=1}^{k} NET_i}{k} \pm 2 \cdot \sqrt{\hat{\sigma}^2(NET)/k}$$

Equivalently, this confidence interval may be computed roughly from the graphs as

$$(7') \qquad CUM_k \pm 2[(CUM_k)(CV_k]$$

Real Trend in NET Over Time. It is possible that superposed onto the artificial trend is a real, biologically based trend in net retention as well. Such a supposition is suggested by the fact that all three of the CUMCa plots show a decrease as d increases, which under an assumption of stationarity in NET (i.e., no trend) would occur only one time in roughly 3 • 3 • 3 = 27 (supposing three informally distinguishable "graph-characters:" increasing, level and decreasing and assuming the other substances' NET's are known stationary; this is obviously only a rough calculation).

By contrast, the apparent trends in the three subjects for N and for K do not appear systematic in any obvious way.

The hypothesis that the CUMCa's exhibit the superposition of a biologically-based trend in NET is supported by the finding that

truncating the first 27 days from the NET Ca's and computing the corresponding "CUM28", i.e.,

$$\text{CUM28 Ca}_d = \frac{\sum_{i=28}^{d} \text{NET Ca}}{d - 27}$$

yields a limiting level to CUM28 Ca significantly lower than that of the corresponding CUM Ca (see Fig. CUM28). This suggests that the subjects had a higher pre-chamber retention (average NET Ca) than that in the in- and post-chamber phases. Biological considerations would lead to an *a priori* expectation in MBE's of the present design that, if any of N, Ca and K exhibit real trend, Ca would be the most likely candidate. This lends support to the foregoing arguments. (Note: Because the pre-chamber period was often exceptional in SMEAT, this truncation technique can often also serve to "flatten out" the artificial trend. We do not discuss details of this use. The degree of "flattening" achieved by truncation varies from plot to plot.)

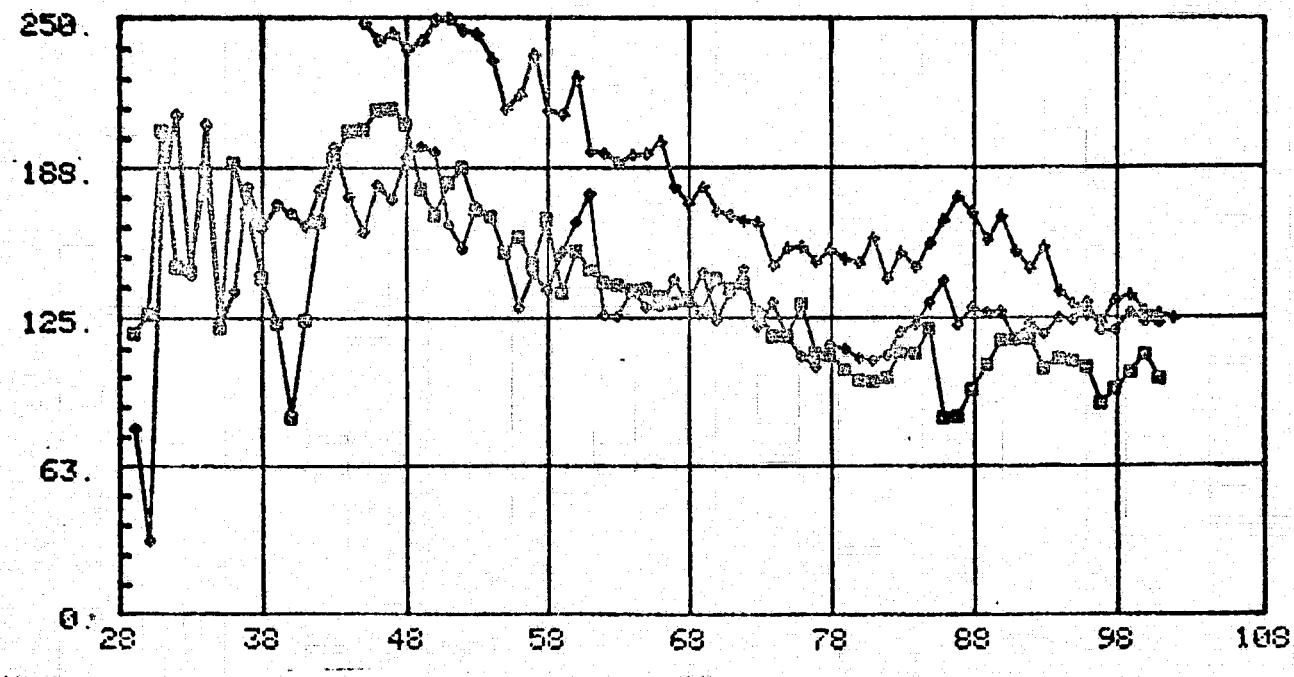



Fig. CUM28. Cumulative calcium retentions for the 3 astronauts over time during the latter part of the study (see text).

Measuring Relative Long-Term Trend. We may obtain quantitative information on the patterns of long-term trends as follows. The preceding discussion suggests the possibility of using the long-term NET K pattern as a "reference" and measuring NET Ca trend with respect to NET K. To do this, we fit the regression models

$$\text{NET Ca} = a + b \cdot \text{NET K} + \text{error}$$

to obtain fitted coefficient $\hat{a}$, $\hat{b}$ and thereby to obtain residuals

$$\text{RES}_i = \text{NET Ca}_i - \hat{a} - \hat{b} \cdot \text{NET K}_i$$

We now regard the RES series as having a 6-day period and perform a two-way analysis of variance on RES and any patterns to the row effects, if "non-random," giving us infomration on the differences between the long-term pattern of Ca retention from that of K. Tables AVres 1-3 show the results for S, C and P, respectively, of the regressions followed by the AV on the residuals (excluding the irregular ends of the MBE's). In all three cases, the row effects clearly exhibit a non-random favoring of positive ones for the early periods of the MBE's and negative row effects for the later periods. This relative trend seems strongest for S. Thus, the astronauts all exhibited systematic decreases in their average Ca retentions as their MBE's progressed, as compared to their K retentions as measured by the residual AV.

Note 1. As is the case with all three substances, any trends in NET K values represent the superposition of biologically based trend and bias due to the MBE procedure. Thus, although the K retentions are relatively stable (see Fig. CUM 12, remembering the end-correction required for SCUMK), their numerical values seem to be even less realistic than those of N and Ca. If we can assume, however, that the bias due to the MBE procedure is constant throughout the range of the MBE considered (in our

Table AVres 1

1: SNETCA = A+B*SNETK

NOB = 90 NOVAR = 2
RANGE = 1 TO 90
RSQ = 0.05747 CRSQ = 0.04675 F(1/88) = 5.365
SER = 447.5840 SSR = 1.763E+07 DW(0) = 1.99

| COEF | VALUE | ST ER | T-STAT |
|---|---|---|---|
| A | 73.42530 | 68.14380 | 1.07751 |
| B | 0.14133 | 0.06102 | 2.31633 |

TWO WAY ANALYSIS OF VARIANCE

| | SSQ | D.F. | MEAN SQUARE | F(RESID) | PROBABILITY |
|---|---|---|---|---|---|
| ROW | 2.833788E+06 | 14 | 202413. | 1.1038 | 0.370066 |
| COL | 1.958827E+06 | 5 | 391765. | 2.13637 | 0.071094 |
| RES | 1.283654E+07 | 70 | 183379. | | |
| TOT | 1.762917E+07 | 89 | | | |

GRAND MEAN = 0.000178

| ROW/COL | ROW EFFECTS | COL EFFECTS |
|---|---|---|
| 1 | 211.501 | -51.4595 |
| 2 | 237.641 | -173.211 |
| 3 | -219.348 | -4.04297 |
| 4 | 311.832 | 278.503 |
| 5 | 163.497 | -122.667 |
| 6 | 47.6747 | 72.8774 |
| 7 | 46.4625 | |
| 8 | 157.085 | |
| 9 | -174.453 | |
| 10 | -120.016 | |
| 11 | -185.987 | |
| 12 | -2.93294 | |
| 13 | -258.587 | |
| 14 | -89.5843 | |
| 15 | -124.785 | |

Table AVres 2

1: CNETCA = A+B*CNETK

NOB = 90 NOVAR = 2
RANGE = 1 TO 90
RSQ = 0.25886 CRSQ = 0.25043 F(1/88) = 30.735
SER = 429.7430 SSR = 1.625E+07 DW(0) = 2.21

| COEF | VALUE | ST ER | T-STAT |
|---|---|---|---|
| A | -52.40340 | 59.43460 | -0.88170 |
| B | 0.37537 | 0.06771 | 5.54396 |

TWO WAY ANALYSIS OF VARIANCE

| | SSQ | D.F. | MEAN SQUARE | F(RESID) | PROBABILITY |
|---|---|---|---|---|---|
| ROW | 1.184795E+06 | 14 | 84628.2 | 0.39759 | 0.97119 |
| COL | 167273. | 5 | 33454.5 | 0.157172 | 0.977156 |
| RES | 1.489972E+07 | 70 | 212853. | | |
| TOT | 1.625179E+07 | 89 | | | |

GRAND MEAN = 0.000135

| ROW/COL | ROW EFFECTS | COL EFFECTS |
|---|---|---|
| 1 | 62.6124 | -25.2339 |
| 2 | 27.3992 | -83.3437 |
| 3 | 110.821 | 47.3411 |
| 4 | 13.7148 | 13.6332 |
| 5 | -77.9852 | 25.8645 |
| 6 | 95.7891 | 21.7388 |
| 7 | 108.896 | |
| 8 | 5.38876 | |
| 9 | -124.364 | |
| 10 | -18.2777 | |
| 11 | -175.534 | |
| 12 | 139.157 | |
| 13 | -187.011 | |
| 14 | 171.008 | |
| 15 | -151.614 | |

Table AVres 3

1: PNETCA = A+B*PNETK

NOB = 90 NOVAR = 2
RANGE = 1 TO 90
RSQ = 0.27994 CRSQ = 0.27176 F(1/88) = 34.212
SER = 327.2900 SSR = 9.426E+06 DW(O) = 2.29

| COEF | VALUE | ST ER | T-STAT |
|---|---|---|---|
| A | -46.16990 | 51.84700 | -0.89050 |
| B | 0.35849 | 0.06129 | 5.84913 |

TWO WAY ANALYSIS OF VARIANCE

| | SSQ | D.F. | MEAN SQUARE | F(RESID) | PROBABILITY |
|---|---|---|---|---|---|
| ROW | 1.050686E+06 | 14 | 75049. | 0.666759 | 0.798261 |
| COL | 496734. | 5 | 99346.8 | 0.882628 | 0.497449 |
| RES | 7.879057E+06 | 70 | 112558. | | |
| TOT | 9.426478E+06 | 89 | | | |

GRAND MEAN = 0.000197

| ROW/COL | ROW EFFECTS | COL EFFECTS |
|---|---|---|
| 1 | -41.7425 | -116.232 |
| 2 | 85.6805 | -45.6011 |
| 3 | -195.755 | 98.4011 |
| 4 | 103.374 | -8.38895 |
| 5 | 239.207 | -15.0503 |
| 6 | 55.6263 | 86.8712 |
| 7 | 36.5577 | |
| 8 | 10.1908 | |
| 9 | 51.8806 | |
| 10 | -27.561 | |
| 11 | -146.935 | |
| 12 | -153.013 | |
| 13 | -62.4802 | |
| 14 | -8.34354 | |
| 15 | 53.3137 | |

instances, days 1 through 90), then it is easy to see that the procedure proposed here is equivalent to comparing NET Ca with the biologically based part of NET K alone, as follows. If we let $\Delta$ equal the constant bias, so that

$$\text{NET } K_i = \text{NET } KB_i + \Delta$$

where $\text{NET } KB_i$ is of course the biologically based portion of NET K on the $i^{th}$ day, then it may be verified that

$$\hat{a}_B = \hat{a} + \hat{b}\,\Delta$$

$$\hat{b}_B = \hat{b}$$

where $\hat{a}_B$, $\hat{b}_B$ are the coefficients in a regression of NET Ca on NET KB. Thus, if

$$\text{RES } B_i \equiv \text{NET } Ca_i - \hat{a}_B - \hat{b}_B \text{ NET } KB_i$$

we have

$$\text{RES B} = \text{NET Ca} - (\hat{a} + \hat{b}\,\Delta) - \hat{b}(\text{NET K} - \Delta) = \text{RES},$$

so that the residual with respect to NET K and with respect to NET KB are the same, as we set out to show.

Role in Future MBE Analysis. This technique, when used in an "updated" manner as an MBE progresses, and making comparisons among several substances at least some of whose trend characters are approximately predictable in advance, may allow inferences about evolving trend in an MBE even before all substances have stabilized, i.e., have zero trend, in their NET's.

We unfortunately must conclude that however accurately metabolic balance studies are conducted they will inevitably lead to uncertain results. The usual experience is that the more nutrient fed the more "apparent retention" is obtained and that the results are not made more interpretable by very long-term studies. Since the reasons for the "errors" are unknown (even though skin losses seem the most logical explanation), comparative studies with different environmental conditions or different subjects must be viewed with great caution and suspicion. Calcium balances are particularly difficult to interpret. Adapation to new conditions may require months and the losses or gains in total body calcium which appear to occur, as in SMEAT, may be insignificant in terms of the total body supply. Clearly, one cannot project losses observed over even a few months to long periods.

Studies utilizing activation analysis for estimation of total body calcium or nitrogen or the measurement of total body potassium by $^{40}K$ counts or the estimation of total potassium pools appear to offer the best direction for research. Even though the accuracy of such measures is difficult to establish, they should be satisfactory for comparative measurements, before and after experimental treatments, as in the exposure to weightless conditions. Until such independent measures of total body composition are available, the interpretation of balance studies will remain questionable.

ACKNOWLEDGEMENTS

The bulk of the machine computation documented in this paper was done through the courtesy of the National Bureau of Economic Research and its TROLL language. The capabilities of TROLL facilitated this research in an indispensable manner.

Dr. Kenneth Wachter of the Harvard Statistics Department made helpful comments regarding the distinction between the restricted and the wider inference and the associated variance estimates.

REFERENCES

1. Arnold, B. Error Analysis: A 90-Day Report. Skylab Food Test and Integration. Technology Inc. Contract NAS-9-11843.

2. Young, V. R., Y. S. M. Taylor, W. M. Rand and N. S. Scrimshaw. (1973) Protein requirements of man: Efficiency of egg protein utilization at maintenance and submaintenance levels in young men. J. Nutr. 103: 1164.

3. Calloway, D. H. and S. Margen. (1971) Variation in endogenous nitrogen excretion and dietary nitrogen utilization as determinants of human protein requirement. J. Nutr. 101: 205.

4. Clark, H. E., J. M. Howe, J. L. Magee and J. L. Malzer. (1972) Nitrogen balance of adult human subjects who consumed four levels of nitrogen from a combination of rice, milk and wheat. J. Nutr. 102: 1647.

5. Clark, H. E., W. H. Moon, J. L. Malzer and R. L. Pang. (1974) Nitrogen retention of young men who consumed between sixteen and eight grams of nitrogen from a combination of rice, wheat, chicken and milk. Amer. J. Clin. Nutr. 27: 1059.

6. King, J. C., D. H. Calloway and S. Margen. (1973) Nitrogen retention, total body $^{40}K$ and weight gain in teenage pregnant girls. J. Nutr. 103: 772.

7. Abernathy, R. P., M. Speirs, R. W. Engel and M. E. Moore. (1966) Effects of several levels of dietary protein and amino acids on nitrogen balance of preadolescent girls. Amer. J. Clin. Nutr. 19: 407.

8. Grindley, H. S. (1912) Studies in nutrition. An investigation of the influence of salt peter on the nutrition and health of men with reference to its occurrence in cured meats. Vol. IV. The experimental data of biochemical investigations. University of Illinois.

9. Mitchell, H. H. and T. S. Hamilton. (1949) The dermal excretion under controlled environmental conditions of nitrogen and minerals in human subjects, with particular reference to calcium and iron. J. Biol. Chem. 178: 345.

10. Calloway, D. H., A. C. F. Odell and S. Margen. (1971) Sweat and miscellaneous nitrogen losses in human balance studies. J. Nutr. 101: 775.

11. Forbes, G. B. (1973) Another source of error in the metabolic balance method. Nutr. Rev. 31: 297.

12. Wallace, W. M. (1959) Nitrogen content of the body and its relation to retention and loss of nitrogen. Fed. Proc. 18: 1125.

13. Duncan, D. L. (1958) The interpretation of studies of calcium and phosphorus balance in ruminants. Nutr. Abstr. Rev. 28: 695.

14. Consolazio, C. F., L. O. Matoush, R. A. Nelson, L. R. Hackler and E. E. Preston. (1962) Relationship between calcium in sweat, calcium balance, and calcium requirements. J. Nutr. 78: 78.

# APPENDIX I

*Periodicity and Correlation Structures in the Data Series*

Motivation. The rather substantial CV's, even the long-term CV's, move us to consider whether any systematic structure within or between the 27 time series for I, U and F summarized in Table 1 in Section I can allow reducing the $\hat{\sigma}^2$ as displayed; Table 1 directs our attention in particular to the urinary nitrogen, urinary potassium and fecal calcium which contribute most to their respective $\hat{\sigma}^2$(NET).

Character of Finding. A naive first hypothesis is that what is excreted in the urine or feces is a function of the dietary input n days before. We have explored the data with such hypotheses in mind, attempting to predict $F_i$ from its precursors, etc.

Our principal finding is that such variance reduction is not available to any significant degree, and in this sense the estimates in Table 1 are realistic characteristics of MBE procedures. We found the interdependencies too weak compared to the inherent variabilities of the system.

A major feature of the data that hindered the present inquiry was the highly irregular quality inherent in the F series. Although U can be presumed to be coming out associated with the I of hours before, F can be delayed by a variable number of days that need bear no relation to the diet cycle.

The hour of last urination before collection, as another data variable, could allow adjusting by some partitioning scheme for the "collection effect;" this could be especially important for N and K and other substances where the primary excretory channel is urine.

Significant Correlations. Table AI.1 summarizes the significant correlations among the input and output series. "VOL" denotes the daily urine volume. (S = SPT, C = CDR, P = PLT in this Appendix) SPTUN (-3) denotes the "3-days lagged" SPTUN series.

The W5-series denotes smoothed fecal series by the triangular moving average,

$$W5F_i = 0.05\ F_{i-2} + 0.25\ F_{i-1} + 0.40\ F_i + 0.25\ F_{i+1} + 0.05\ F_{i+2}$$

We do not find any significant correlations for any of the three astronauts of the fecal series, thus smoothed, with either the input or the urine series, although there seems to be a trace of covariation with some lagged values of the variable,

$$GUT \equiv I - U$$

which measures how much of the diet input is not excreted by the urine on a given day.

The less-than full ranges in Table AI.1 were chosen so as to exclude clearly aberrant end-values of the MBE's.

The high correlations with VOL are somewhat surprising, but this is not a useful issue for us to discuss so we are simply documenting the results.

The variables with the "SQ(ueezed)"-prefix denote the fecal series with the 0's deleted; these are the appropriate measures of association of substances' fecal measurements, because the exact-zero entries inflate the correlation "artificially." The high contemporaneous correlations among the SQF's are not surprising as the feces can be expected to be fairly homogeneous on a macroscopic scale with respect to the several substances.

TABLE AI.1

Significant Correlations

| Variable | Variable | Range | Correlation |
|---|---|---|---|
| SPTFCa | SPTFK | 25 - 78 | 0.851 |
| | | 1 - 101 | 0.791 |
| SPTUCa | SPTVOL | 1 - 101 | 0.937 |
| SPTUN | SPTVOL | 18 - 83 | - 0.460 |
| SPTUCA | SPTUN(-3) | 20 - 83 | 0.223 |
| SPTUN | SPTUN(-1) | 20 - 83 | 0.203 |
| SPTUCa | SPTUK | 20 - 83 | 0.512 |
| SPTUCa | SPTUK(-1) | 20 - 83 | - 0.239 |
| CDRUCa | CDRUCa(-2) | 6 - 102 | 0.261 |
| CDRUCa | CDRUCa(-4) | 6 - 102 | 0.215 |
| SPTUCa | SPTUK | 25 - 78 | 0.443 |
| CDRFCa | CDRFK | 1 - 102 | 0.952 |
| SQCFCa | SQCFK | 1 - 64 | 0.899 |
| SQPFCa | SQPFK | 1 - 87 | 0.838 |
| SQSFCa | SQCFCa(-4) | 6 - 84 | 0.255 |
| SQPFCa | SQPFCa(-2) | 6 - 87 | 0.270 |
| | (-3) | | 0.279 |
| SQSFCa | SQSFK | 1 - 84 | 0.716 |
| | | 15 - 70 | 0.754 |
| CDRFN | CDRFCa | 1 - 70 | 0.922 |
| CDRUCa(+3) | CDRUCa(-3) | 24 - 80 | 0.409 |
| | (-2) | | - 0.376 |
| PLTUCa(+3) | PLTUCa | 4 - 86 | - 0.311 |
| | (-3) | | 0.257 |
| SPTUCa | SPTVOL | 1 - 101 | 0.626 |
| CDRUN | CDRVOL | 1 - 102 | 0.229 |
| CDRUCa | | | 0.539 |
| PLTUN | PLTVOL | | 0.371 |
| PLTUCa | | | 0.373 |
| SPTUK | SPTVOL | 1 - 101 | 0.451 |
| CDRUK | CDRVOL | 1 - 102 | 0.579 |
| PLTUK | PLTVOL | | 0.450 |
| SGUTCa | W5SFCa(+2) | 3 - 95 | 0.221 |
| CGUTCa | W5CFCa(+1) | 3 - 96 | 0.114 |
| PGUTCa | W5PFCa | | 0.230 |
| | (+1) | | 0.114 |
| | (+4) | 3 - 96 | 0.149 |

The highly varying non-gaussian character of the fecal series, coupled with the sense that $F_i$ represents the I from a variable number of days before, suggests that some non-parametric technique might be helpful.

We have considered the series

$$FRl_i = \text{rank } (F_i) \text{ in the set } \{F_j\}_{j=1}^{\text{maxium day no.}}$$

where all zeros are given a rank of 1, so that for example

$$SFCa_{20} - SFCa_{35}$$

shown in Table raw, are replaced by their ranks as in Table ranks. The correlation matrix of the rank-series SFCaRl is shown in Table corr; it is suggestive but again the magnitudes are too small to be useful.

Spectral Analyses--Background and An Example. A more delicate technique than autocorrelations for distinguishing periodicities in a time series is spectral analysis. Spectral analysis regards a time series as being the superposition of sinusoidal waves over a continuum of frequencies; a graph of the "spectrum" versus frequency (in cycles/day) exhibits the magnitude of the contribution of these harmonics at each frequency to the total variance of the series. Fig. 2 is the (logarithm of) the spectrum versus frequency (which we refer to as the "spectral plot") for a series with a known strong, but not perfect, periodic component with

$$\text{frequency} = \frac{1 \text{ (diet) cycle}}{6 \text{ days}} = 0.167, \text{ namely}$$

the series SPTIN shown in Fig. 1. (We exclude the clearly irregular final two periods from the analysis.)

Notice the peak centered at about 0.165, reflecting the diet periodicity of 6 days. Notice further that the spectral plot "tunes in" to a 3-day cycle as well (1 cycle/3 days = 0.333) as shown by the peak centered at about 0.33; a closer examination of Fig. 3 reveals this more subtle feature which was an unintentional feature of the

Table raw

SPTFCA -

| | | | | |
|---|---|---|---|---|
| 20 | 0. | 0. | 0. | 143.92 |
| 24 | 684.42 | 250.4 | 91.6 | 344.47 |
| 28 | 234.25 | 920.64 | 692.37 | 282.45 |
| 32 | 813.12 | 589.92 | 310.84 | 1989.27 |

Table ranks

SFCAR1. -

| | | | | |
|---|---|---|---|---|
| 20 | 1. | 1. | 1. | 5. |
| 24 | 56. | 9. | 3. | 15. |
| 28 | 12. | 67. | 42. | 11. |
| 32 | 59. | 36. | 14. | 75. |

Table corr

TROLL COMMAND: do correl(sfcar1.,sfcar1.(-1),sfcar1.(-2),sfcar1.(-3),sfcar1.(-4
CONTINUE EQUATION: sfcar1.(-6));

RANGE 7 101 CORRELATION MATRIX

| | SFCAR1. | ARG2 | ARG3 | ARG4 | ARG5 | ARG6 | ARG7 |
|---|---|---|---|---|---|---|---|
| SFCAR1 | 1.000 | | | | | | |
| ARG2 | 0.024 | 1.000 | | | | | |
| ARG3 | 0.167 | 0.042 | 1.000 | | | | |
| ARG4 | -0.124 | 0.172 | 0.057 | 1.000 | | | |
| ARG5 | 0.015 | -0.117 | 0.167 | 0.038 | 1.000 | | |
| ARG6 | 0.059 | -0.022 | -0.090 | 0.135 | 0.057 | 1.000 | |
| ARG7 | 0.050 | 0.027 | -0.003 | -0.069 | 0.211 | 0.045 | 1.000 |

experimental design.

The sensitivity of the spectral plot to periodicity is illustrated by Fig. 3 which shows SPTIN plotted against SPTIN(-3); if the 3-day cycle were very strong, this plot would appear as a very flat elliptical point-cloud--but in fact it departs little from circularity (the stray values come from the end of the experiment which was excluded from the spectral analysis). Yet the spectral plot tunes in to this weak periodicity with period of 3 days.

For those readers familiar with the technical aspects of computing spectral plots, we include with each of our spectral plots some documentation regarding the particular options used in each case: type of prewhitening (if any), parameter for the "fast Fournier transform" (FFTPARAMETER), etc., as well as the series' variance in principle, but sometimes differing a small amount for numerical reasons). Further details of the algorithms used to produce these plots are available on request.

A Remark on MBE Design. In an MBE with deliberately highly varying dietary contributions from a particular mineral, it is important to investigate before the experiment the harmonic structure within the diet cycle, lest information be obscured through having periodicities with cycle-lengths that are multiples of each other, making it impossible to distinguish which proposed effect, in fact, drives the output. It might, in fact, be useful to investigate an MBE which was deliberately designed to allow variability in intake in order to more fully examine these relationships.

Specific Findings. Fig. 4 shows SPTUN rising slightly in the final periods; its spectrum (in Fig. 5, over the full range of SPTUN (1-101) unfortunately gives no support to the hypothesis that the urine recapitulates the cyclicalities of the input. The dietary loadings were apparently too nearly constant to hope for that. Not surprisingly, the more irregular SPTFN (Fig. 5a) has a spectral plot of inconclusive shape as well. The situation is the same in character for CDR's and PLT's U and F plots as well. The SPTICa series, as might be anticipated because they both arise from the same diet, looks the same in character as the SPTIN series. The other diet (I) series does also.

U Spectral Plots. The spectrum of SPTUCa exhibits peaks at frequencies without any obvious interpretation (Fig. 6) in the context of the experiment; CDRUCa's spectrum (Fig. 7) shows a 2-day peak and does seem to recaptulate weakly the 3- and 6-day cycles in CDRICa. PLTUCa's spectrum (Fig. 8) seems to share the same general character although here periods seem closer to 3½ and 7 days (peaks at 0.29 and 0.14)--suggesting that effects with a weekly or biweekly period crept into the experimental procedure. While we do not have any facts about the experiment to settle this issue, it would not be unusual for this to have occurred.

We include the spectral plots of the three astronaut's daily urine volumes because these plots are surprisingly well defined in character; they apparently reflect a sensitivity to the input cyclicality, some biological rhythms, and artifacts of the experimental collection procedures in ways that might warrant further study, although such a treatment would be a rather technical digression for the present paper, involving exploring, e.g., the interpretations in the experimental context of the possible "aliases" of the peak-frequencies. SPTVOL (shown in Fig. 9) while apparently random to the unaided eye, turns up a well defined 6-day peak in its spectrum (Fig. 10). CDRUCa's spectrum (Fig. 11) shows a clear 3-day peak; PLTUCa's spectrum (Fig. 12) shows peaks at 6-7, 4, and 3 days per cycle.

Urinary Ca, Urinary K Cross-Spectral Analysis. In a cross-spectral analysis of SPTUCa and SPTUK, the "coherence", which measures the degree to which the two series are "in step" at different frequencies, turned out as in Fig. 13--several peaks with no obvious interpretation and a peak at nearly a 2-day

cycle. This may reflect either shared biological rhythms, or an artifact of the collection procedure (because what was in the bladder at collection time "will be counted tomorrow even though it belongs to today"), or both. Fig. 14 shows the "gain of SPTUCa to SPTUK, which may be thought of as a regression coefficient of the periodic component in SPTUK on that with the same frequency in SPTUCa. The gain plot shows SPTUCa's harmonics to be poor predictors, i.e., weakly correlated) with the corresponding harmonics of SPTUK--except at a biweekly cycle ($0.28 \approx 2/7$): tracking down this artifact to its source in the experimental procedure could lead to an improvement in accuracy. In this way spectral analysis, because of its sensitivity, can serve as a "troubleshooter" in the experimental procedure, i.e., as an exploratory diagnostic rather than its usual role as a tool of confirmatory data analysis.

Fecal Series. Fig. 15 shows SPTFCa with its highly variable entries caused in part by the apparently random occurrences of zeros (arising on days with no bowel movements). Fig. 16 shows SPTFCa's spectral plot (days 13 to 95); the peak at frequencies within 0.40-0.50 are probably generated by experimental artifacts. There seems to be a 6-7 day peak and perhaps a 3-day peak also; but these patterns are not articulated clearly in the spectra of PLTFCa or CDRFCa, nor are they strong enough to produce significantly non-zero autocorrelations of the corresponding orders; thus any periodic components to the FCA series seems too weak to exploit for any variance-reduction purposes. It will be recalled from Table 1 of the text that the F series of primary concern are those for Ca, for which the primary excretory channel is the feces

rather than the urine, and hence

$$\sigma^2(F) > > \sigma^2(U)$$

Examination of the F series resulting from omitting all zeros, i.e., making the resulting series shorter by the number of zero days, including ANOVA's with appropriately shortened periods, e.g., the shortened SPTFCa series is 84 days in length, and $\frac{84}{101} \approx \frac{5}{6}$ implying an adjusted period of 5 days), does not turn up any stronger patterns.

Fecal-Urine Relationships. The character of our findings relating the F and U series is the same as our intra-F and intra-U findings; weak patterns and correlations can be ferreted out but not strong enough one's exist to allow any significant variance reductions.

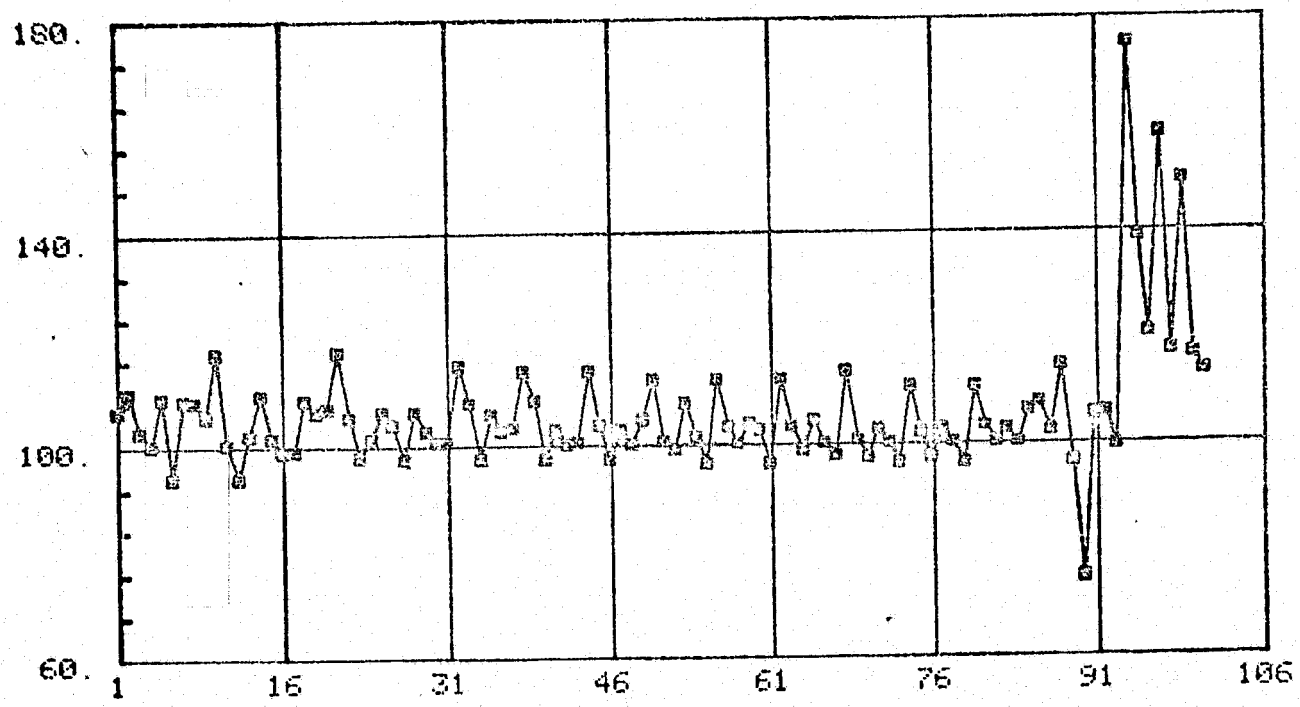
A cross-spectral analysis of SPTFCa and SPTUCa, for example, shows that their 6-day cycles exhibit a trace of being "in step," but are overall undistinctive in character (Fig. 17). The same applies to the 3 and the 5-day cycles of SPTFK and SPTUK (Fig. 18). The peak at 0.45 cycle/day in these two coherence plots recurs in several other plots, suggesting that one of its aliases (n + 0.45, n + 0.55, $n \geq 0$, an integer) may correspond to an experimental artifact that should be tracked down. For example, the hypothesis that this peak represents an effect caused by the length of the urine collection cycle would lead to computing the quantities:

$$\frac{24}{n + 0.45}, \quad \frac{24}{n + 0.55} \qquad \frac{\text{hours}}{\text{cycle}}$$

n an integer $\geq 0$. The most reasonable such values would seem to correspond to n = 2, approximately a 10 hour cycle. We do not have the requisite information to settle the matter but wish to call attention again to the ways that spectral and cross-spectral analyses can serve as exploratory tools in MBE design.

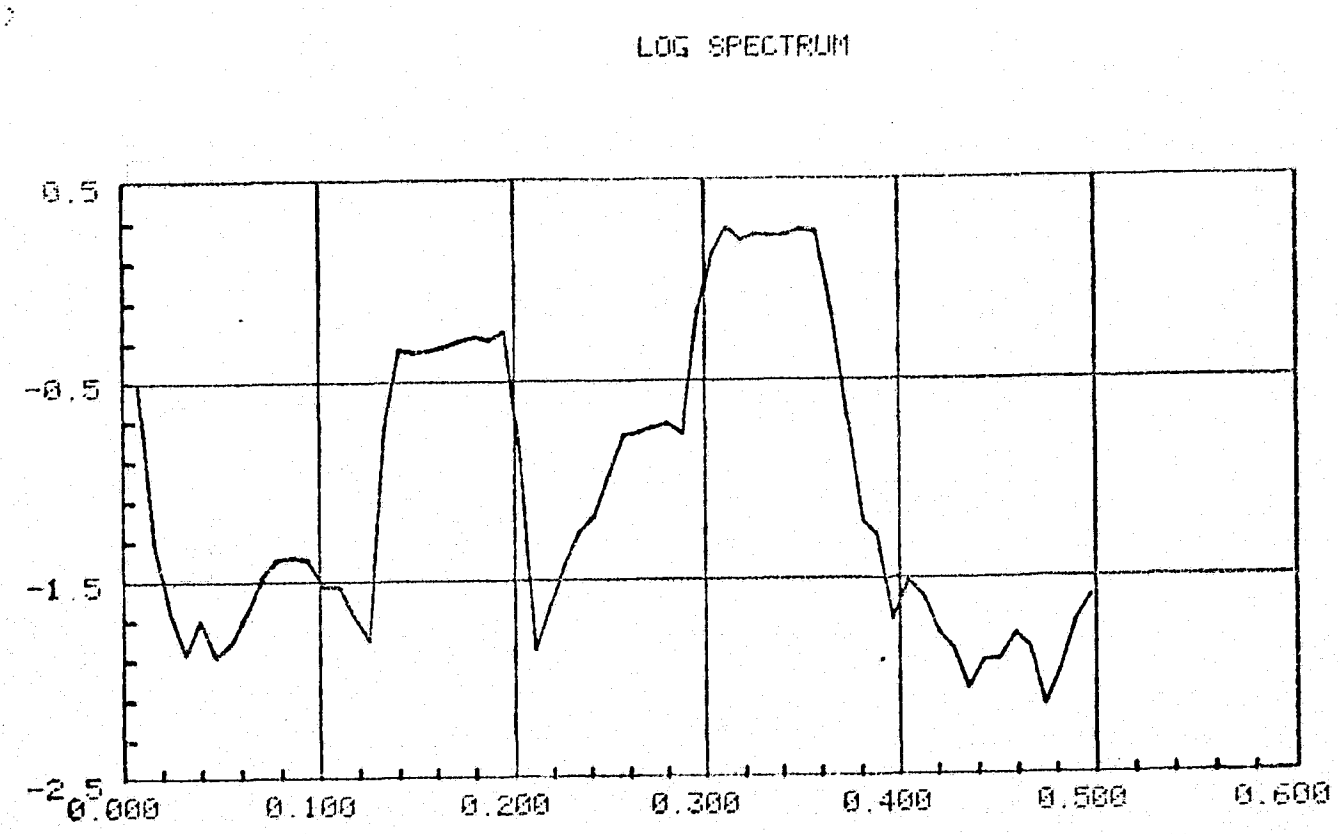
Fig. 19 shows the coherence of the F and U series for CRD's Ca excretion. Notice again the peak at 0.44, the weak 5-6 day peak, the 2-day peak (probably partly an artifact both of the analysis and also of the experimental collection times being at fixed times), and the peak at about 0.11, any of a number of whose aliases could arise for a variety of reasons.

FIG. 1




SYMBOL SCALE NAME
■ #1 SPTIN

FIG. 2




: 8
TYPE OF PREWHITENING: NONE
FILTPARAMETER: 1
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 86
PASIS: 128
BANDWIDTH: 0.082
SPTIN VARIANCE: 28.924 SPECTRUM TOTAL: 28.924

FIG. 3

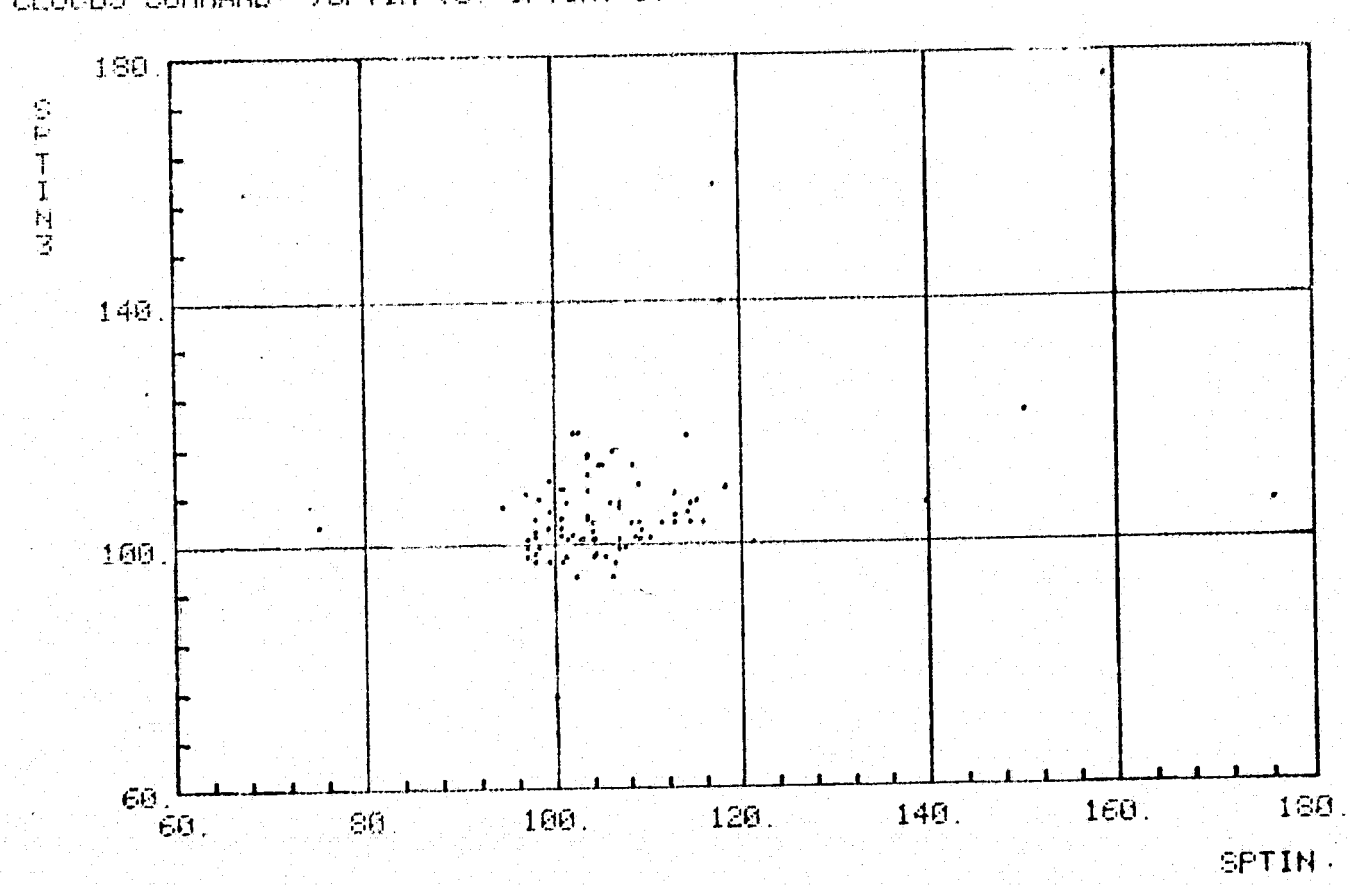
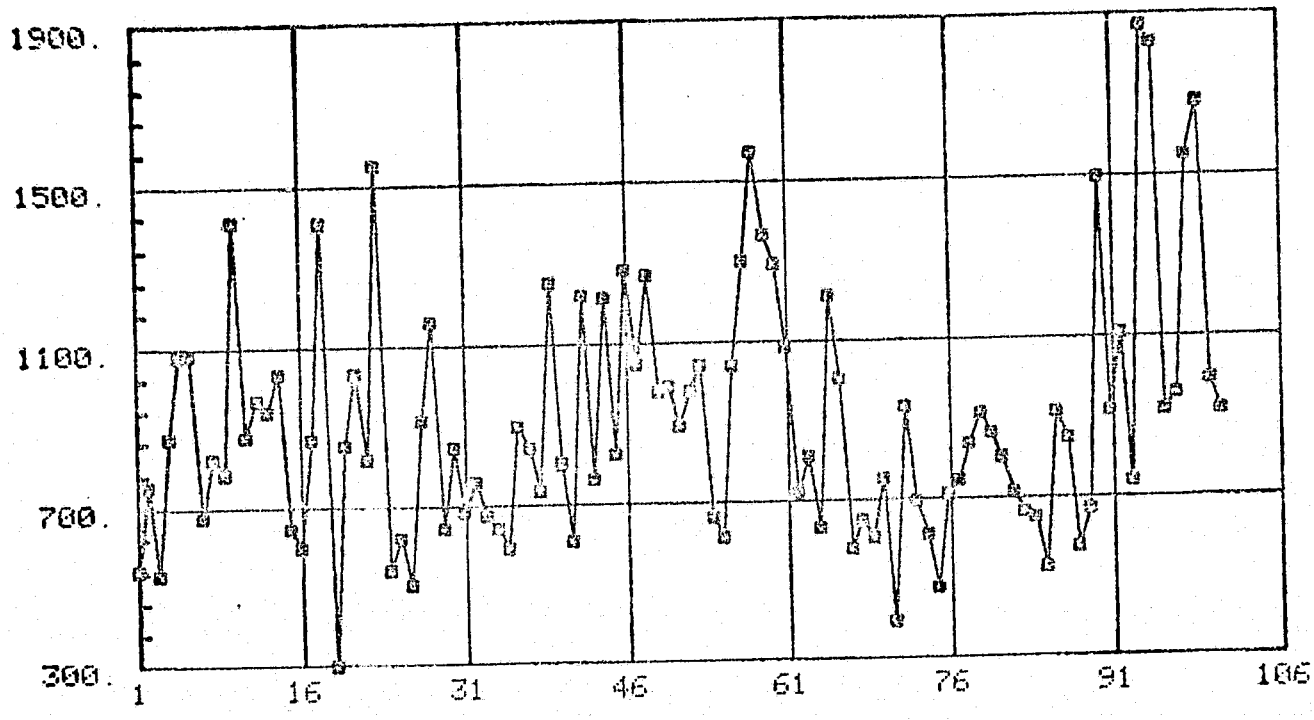
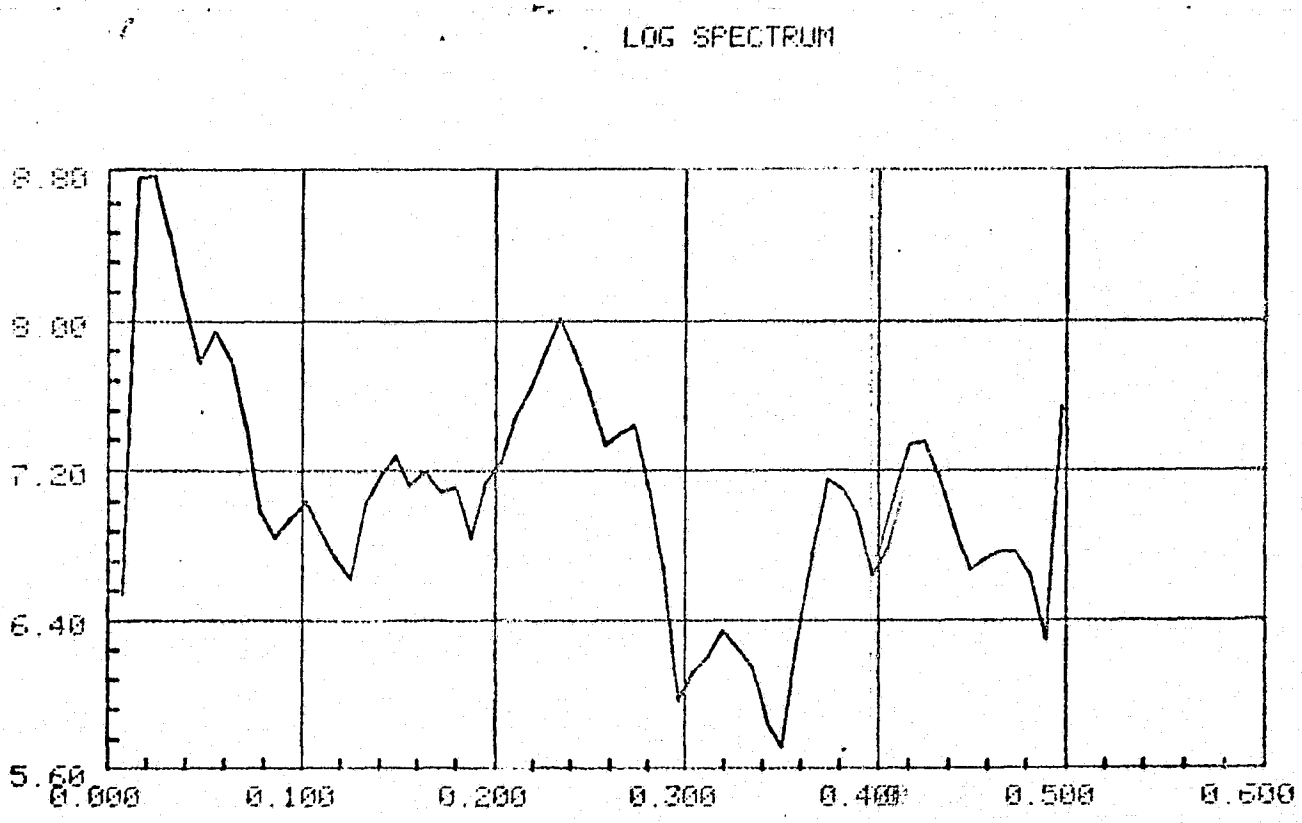

CLOUDS COMMAND: >SPTIN VS. SPTIN)-3(
180.
140.
100.
60.
SPTIN
60.
80.
100.
120.
140.
160.
180.
SPTIN

FIG. 4




SYMBOL SCALE NAME
▫ #1 SPTUN

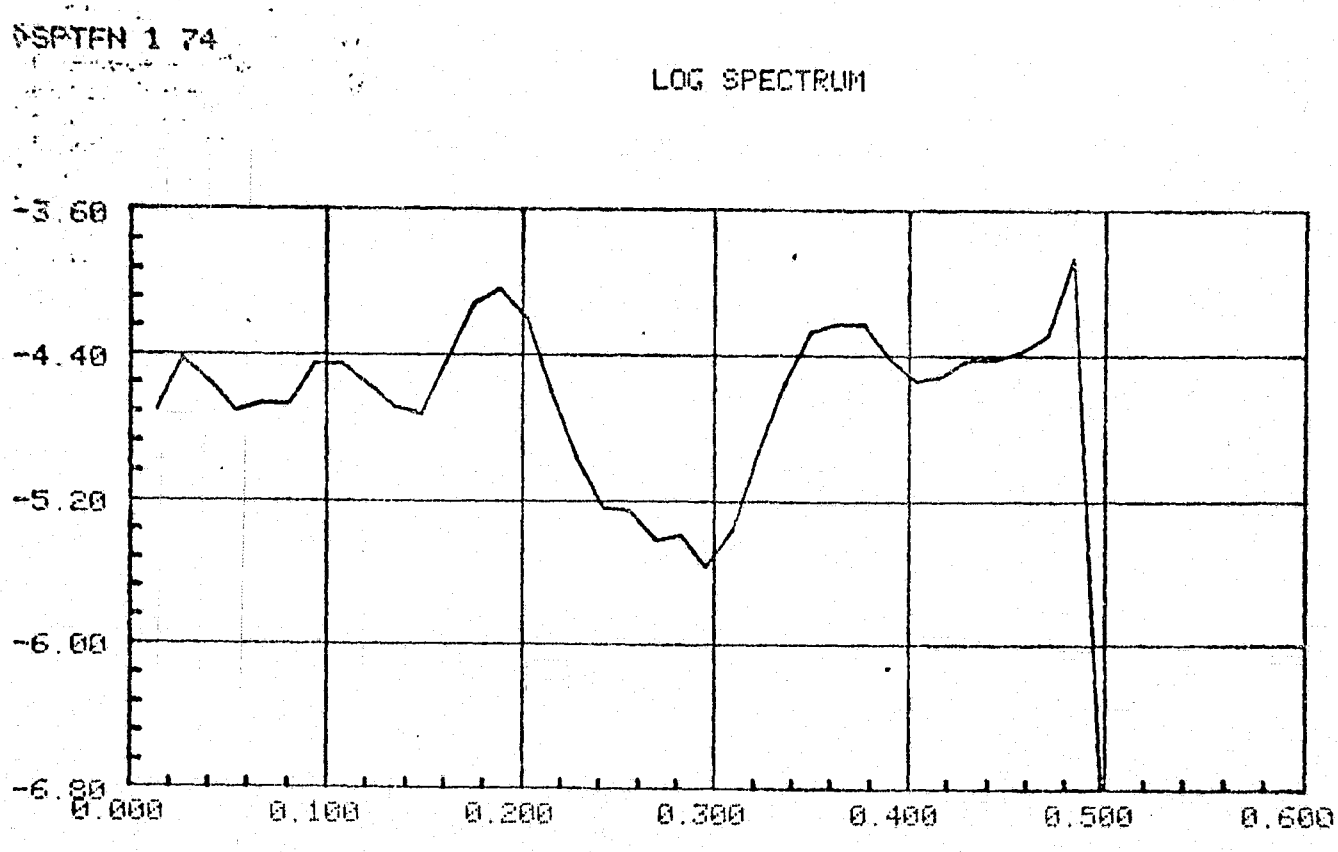
FIG. 5




```
;
TROLL COMMAND: >&SPECT SPTUN WIDTH 5 TRIANG;
%


SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR   RANGE: 4
TYPE OF PREWHITENING: NONE
FFTPARAMETER: 1
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 101
BASIS: 128
BANDWIDTH: 0.043
SPTUN VARIANCE: 96215.700  SPECTRUM TOTAL: 96216.100
```

FIG. 5a.

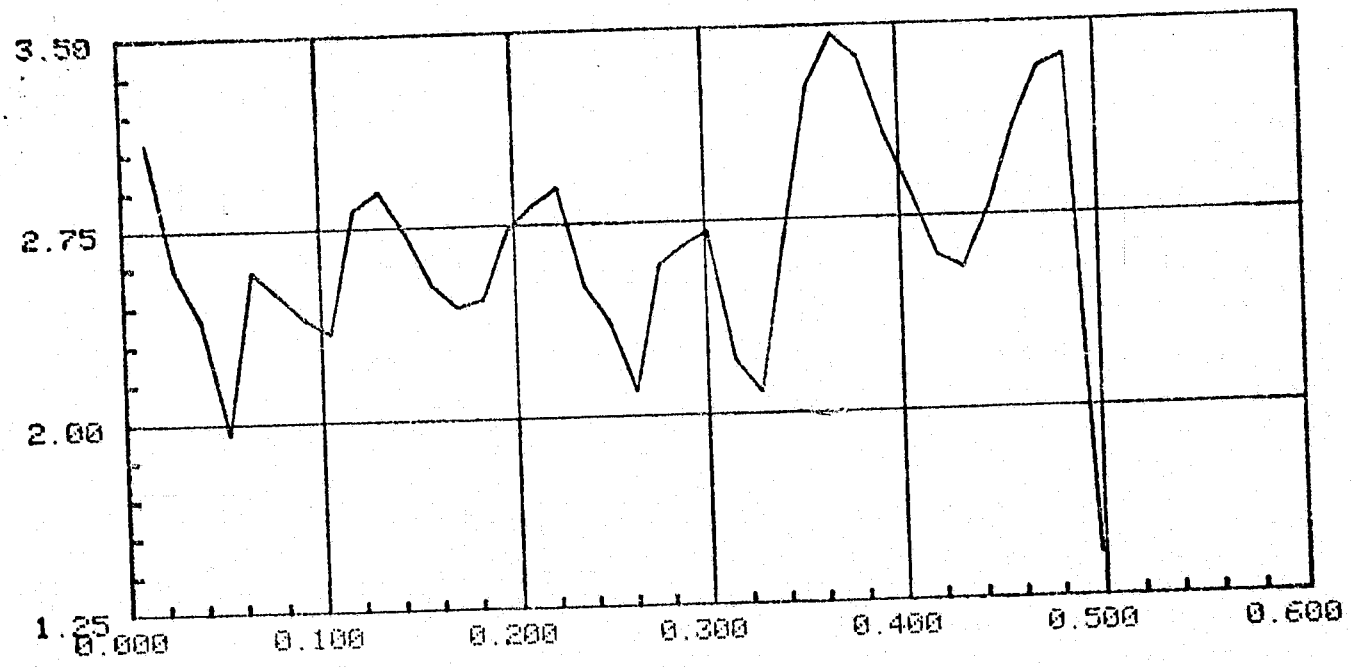



SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 74
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 74
BASIS: 74
BANDWIDTH: 0.071
SPTFN VARIANCE: 0.377 SPECTRUM TOTAL: 0.377

FIG. 6

>SPTUCA 10 TO 85

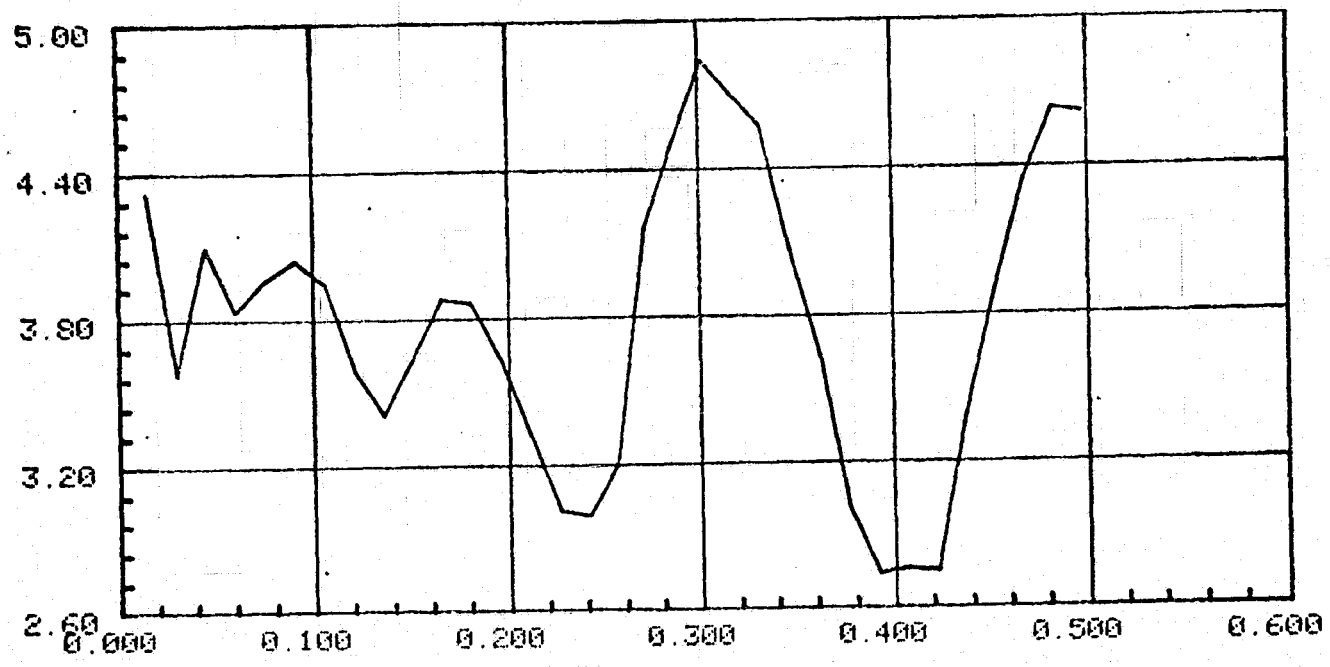
LOG SPECTRUM




X: FREQNCY Y: LSPECT

SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 76
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 76
BASIS: 76
BANDWIDTH: 0.069
SPTUCA VARIANCE: 575.035 SPECTRUM TOTAL: 575.033

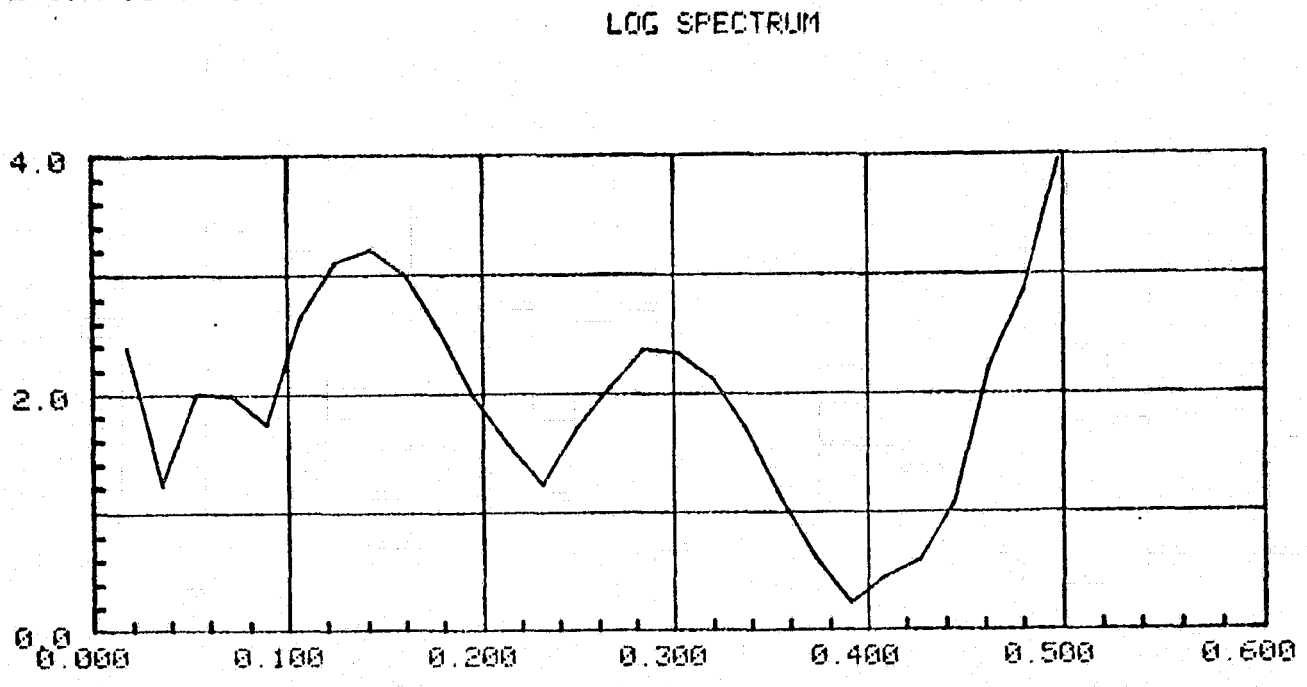
FIG. 7




SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 67
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 67
BASIS: 67
BANDWIDTH: 0.079
CDRUCA VARIANCE: 1594.840 SPECTRUM TOTAL: 1594.840

FIG. 8

```
>PLTUCA 31 TO 86
```

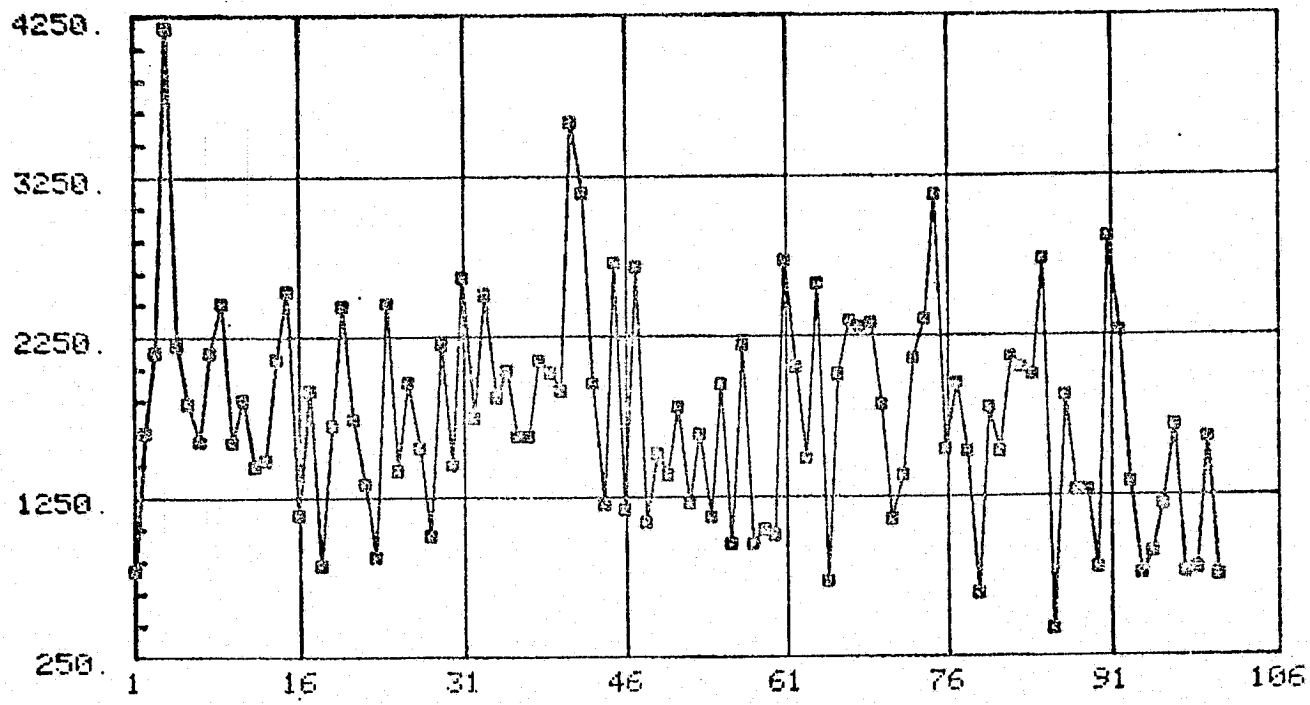



X: FREQNCY-FREQNCY = FREQ(S...SP)

Y: LSPECT-LSPECT = CONSPECT(S...SP,0)

SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 56
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 56
BASIS: 56
BANDWIDTH: 0.094
PLTUCA VARIANCE: 245.345 SPECTRUM TOTAL: 245.346

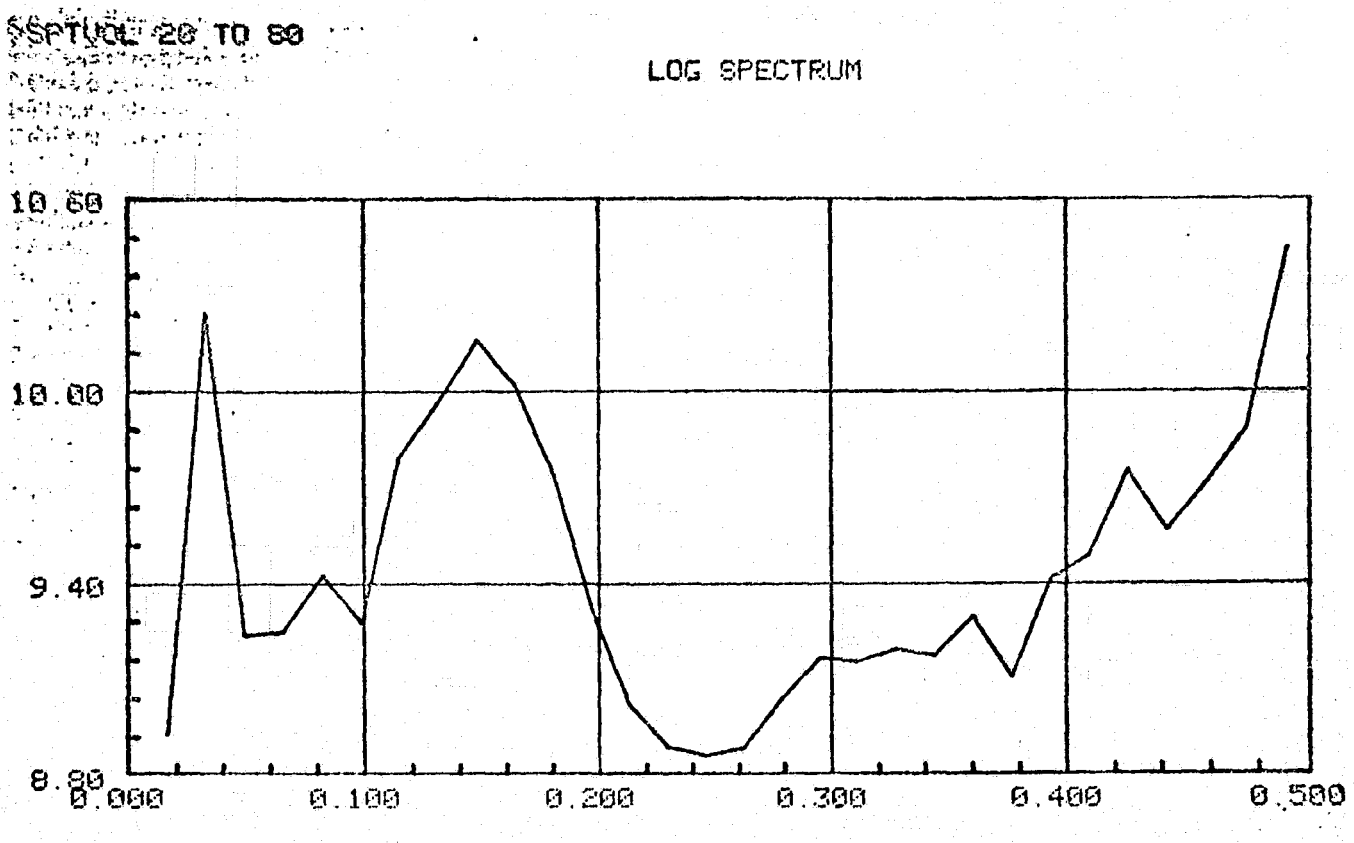
FIG. 9




SYMBOL SCALE NAME

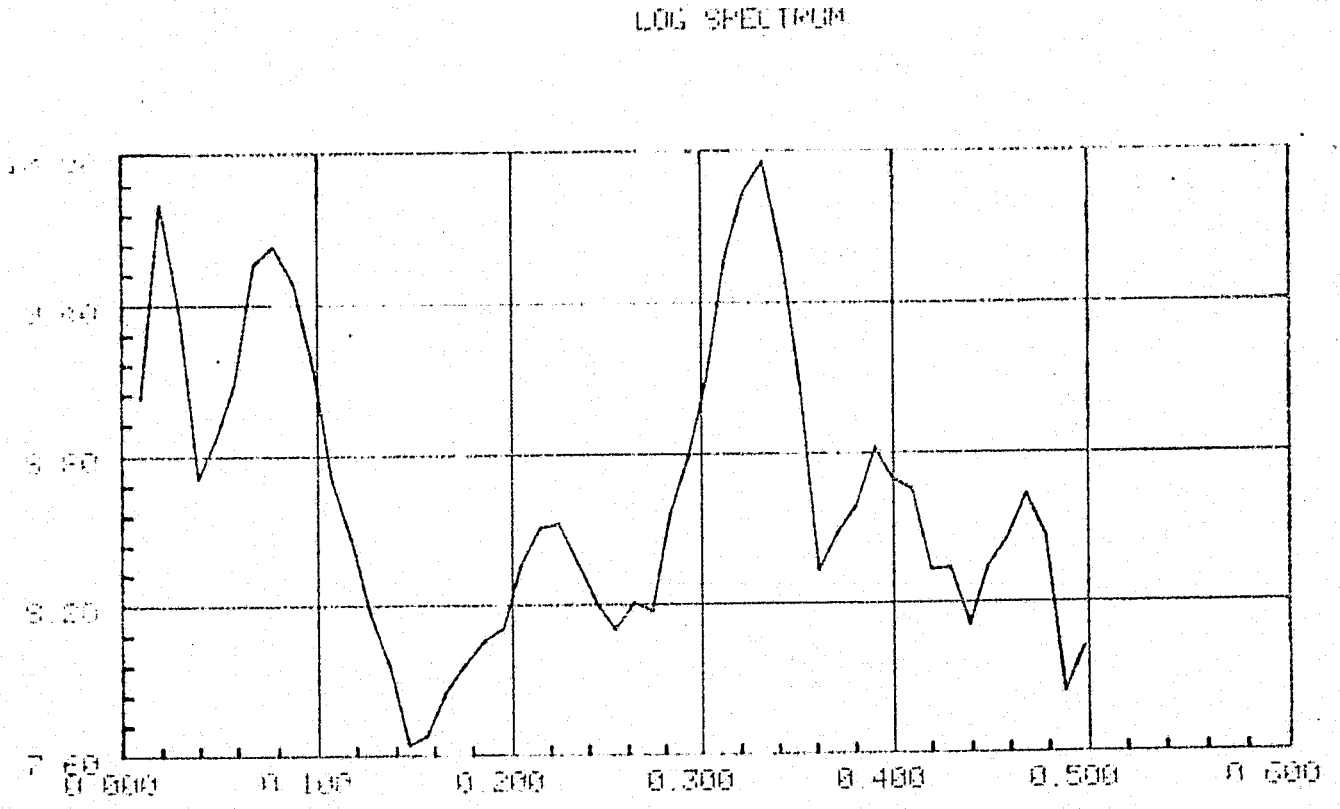
■ #1 SPTVOL

FIG. 10




SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 61
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 61
BASIS: 61
BANDWIDTH: 0.086
SPTUOL VARIANCE: 403509.000 SPECTRUM TOTAL: 403507.000

FIG. 11

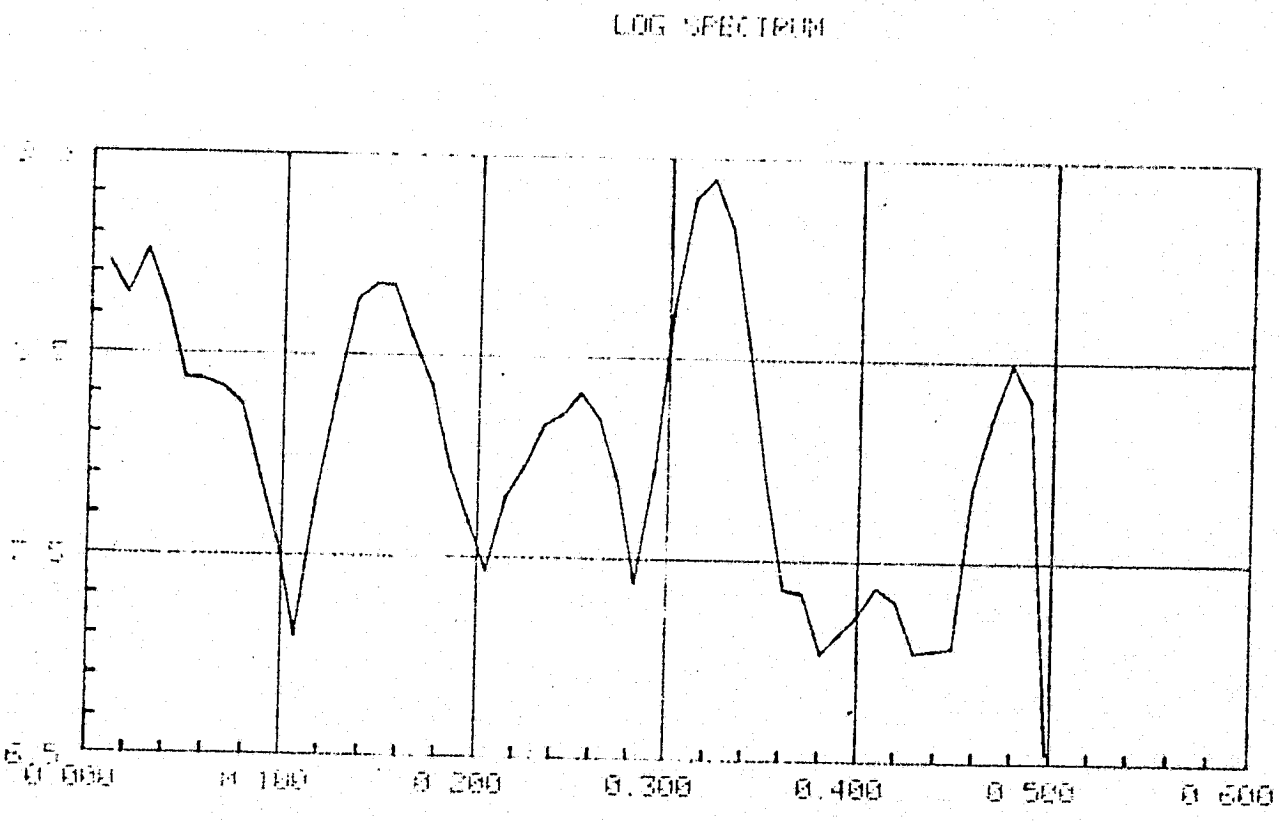



X: FREQNCY Y: LOSPECT



[illegible] STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0 80
AR PARAMETER: 102
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 102
N.LAG 102
BANDWIDTH: 0 052
[illegible] L VARIANCE 341817 000 SPECTRUM TOTAL: 341816 000

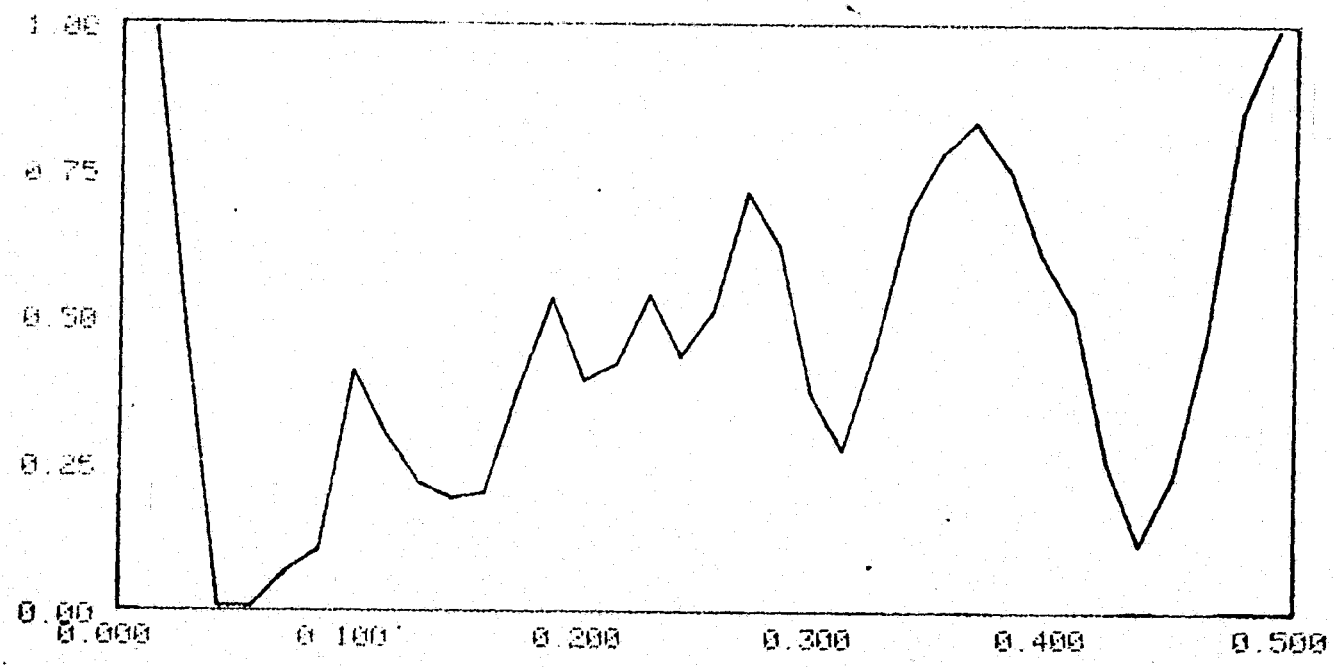
FIG. 12




X: FREQNCY Y: LOGSPCTRM

SPECTRUM STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
PARAMETER: 102
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 102
N: 102
BANDWIDTH: 0.052
TOTAL VARIANCE: 194995.000 SPECTRUM TOTAL: 194992.000

FIG. 13






SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 71
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 71
BASIS: 71
BANDWIDTH: 0.074
NSALIGNMENT
CRITICAL COHERENCE SQUARED (AT 5 PERCENT LEVEL): 0.845
SPTUCA VARIANCE: 602.802 SPECTRUM TOTAL: 602.803
SPTUN VARIANCE: 424355.000 SPECTRUM TOTAL: 424351.000

FIG. 14

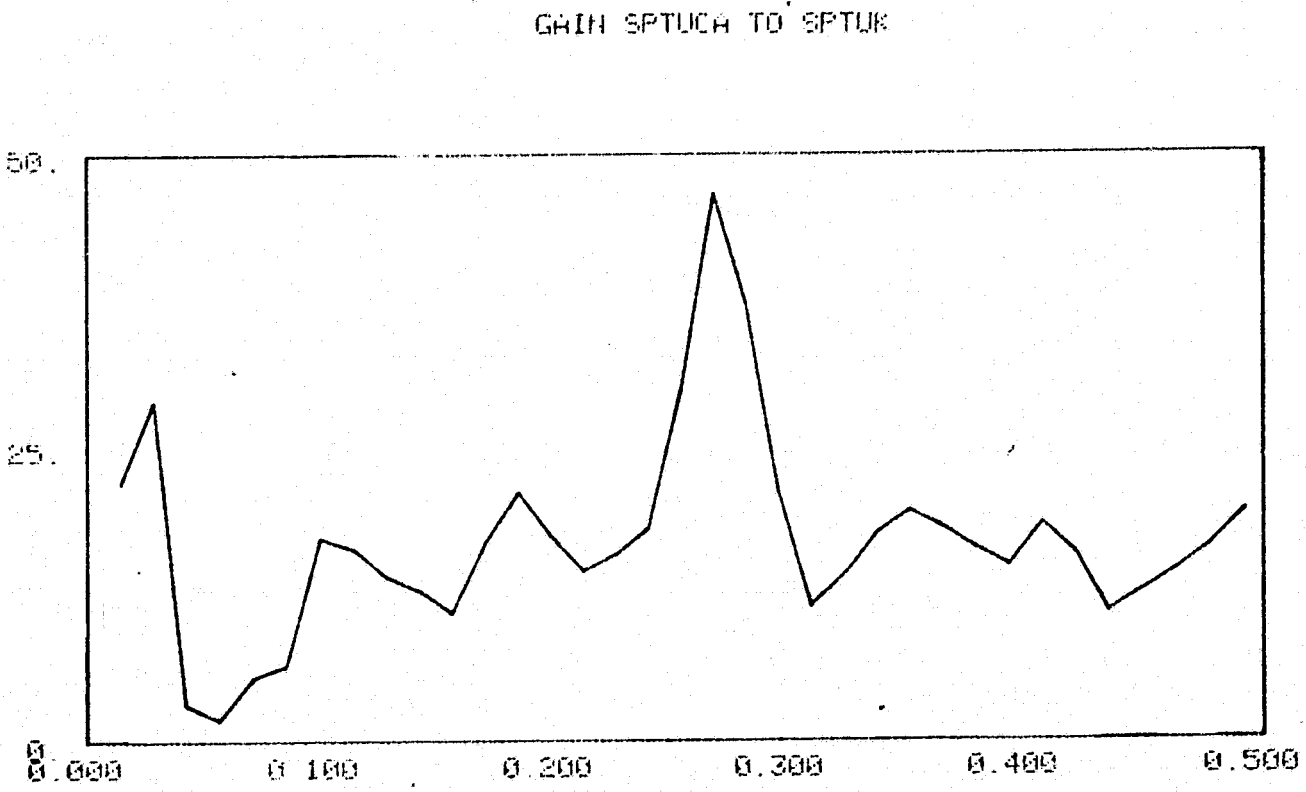

GAIN SPTUCH TO SPTUK
50.
25.
0.000
0.100
0.200
0.300
0.400
0.500
X: FREQNCY Y GAINI2

FIG. 15

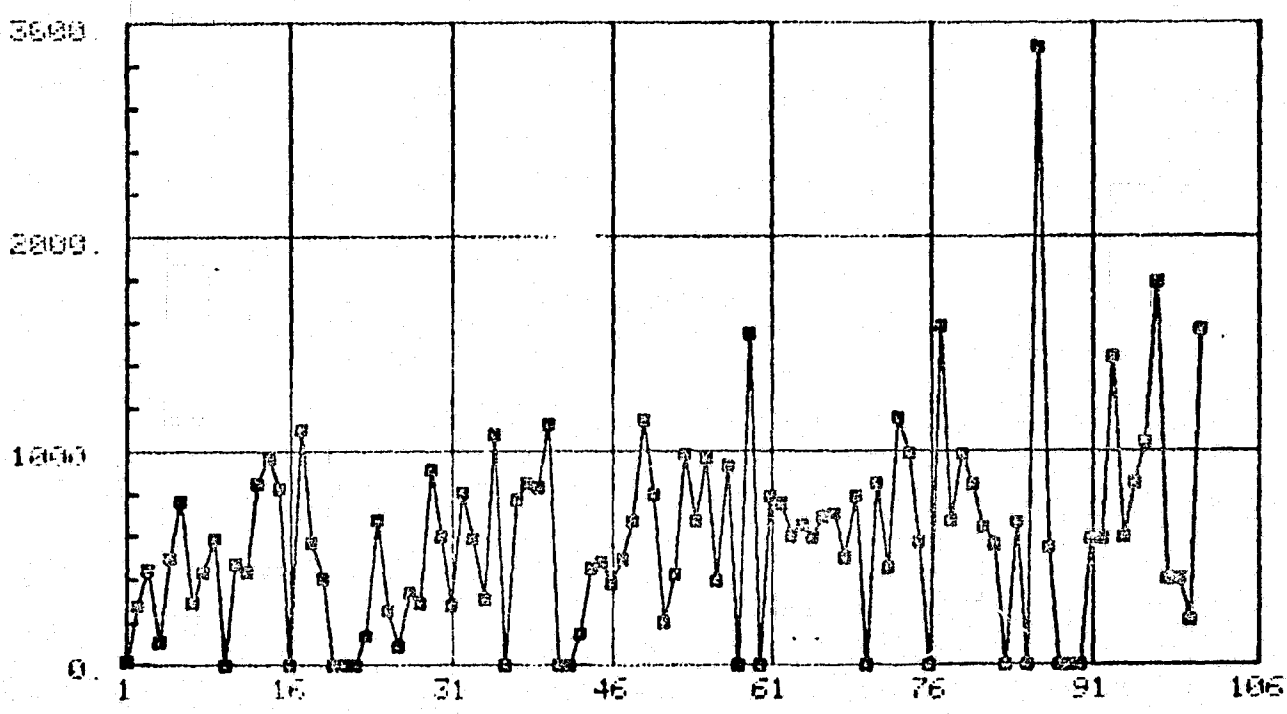
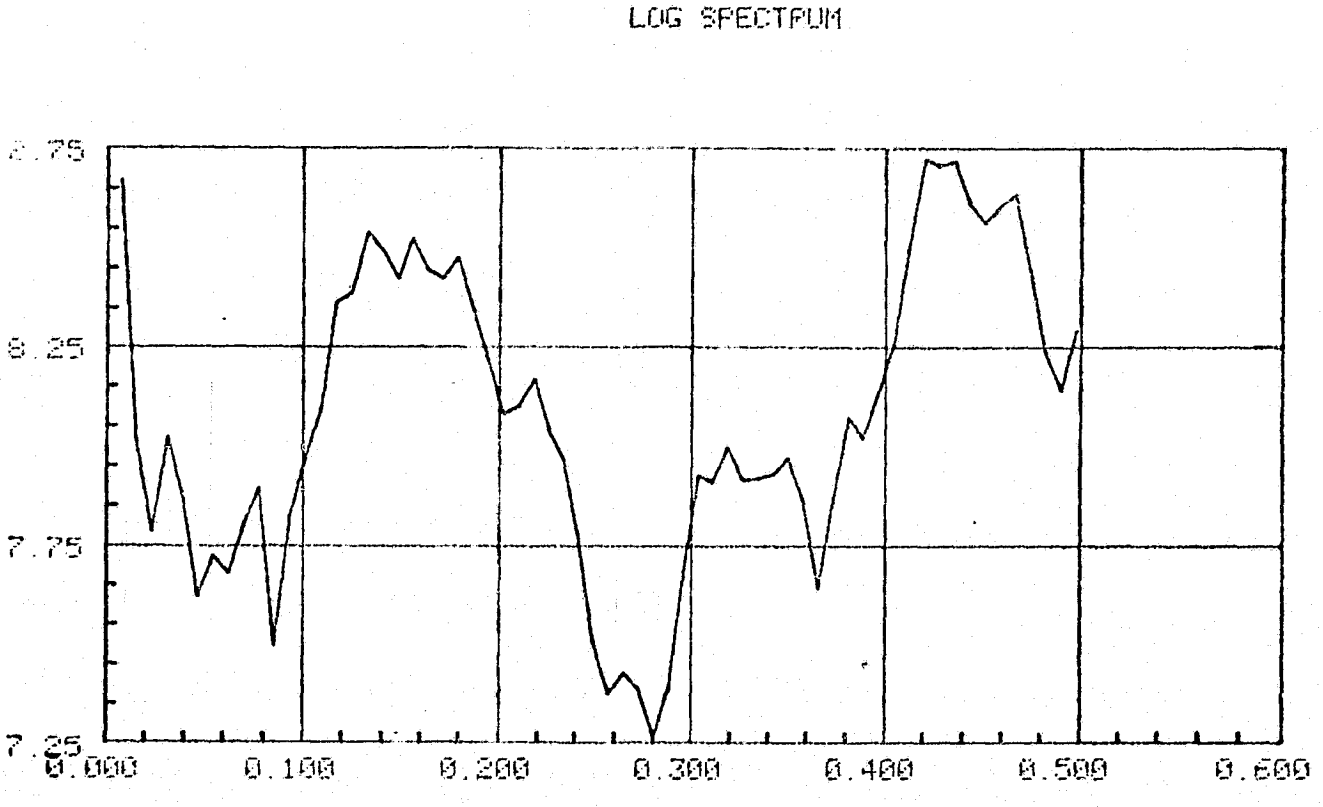
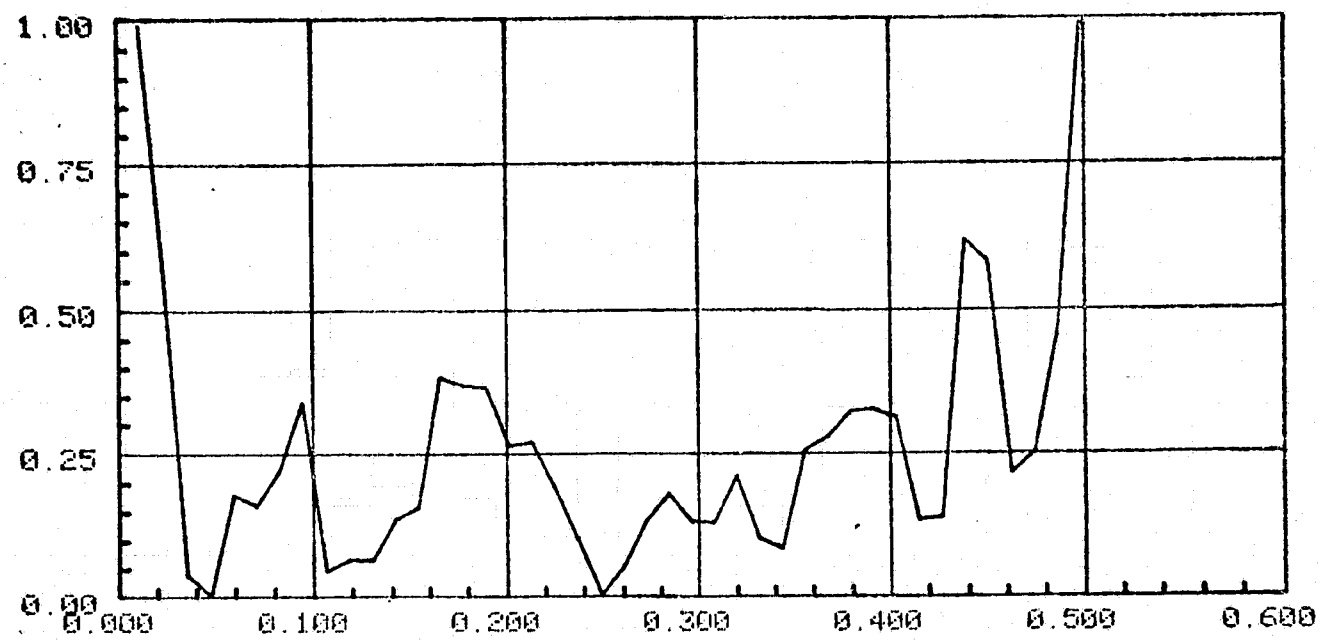

3600.
2800.
1000
0.
1
16
31
46
61
76
91
106
SYMBOL SCALE NAME
#1 SPTFCA

FIG. 16



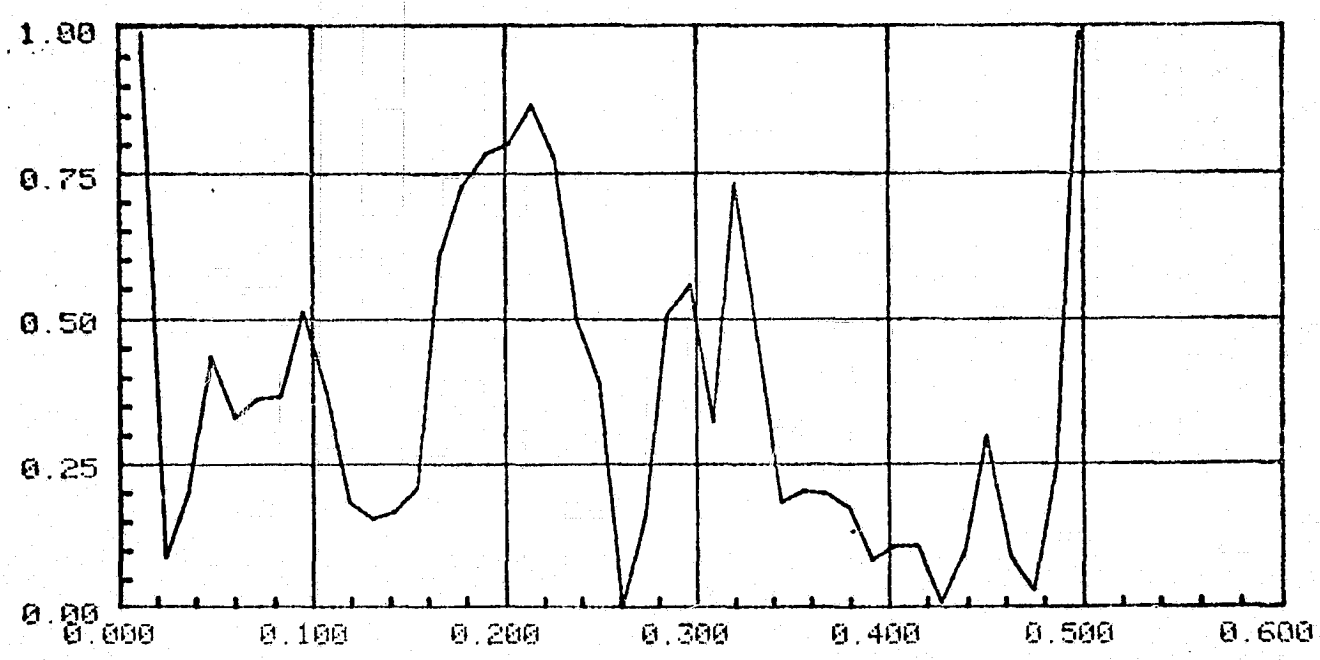

SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: RECTANGULAR RANGE: 8
TYPE OF PREWHITENING: NONE
FFTPARAMETER: 1
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 83
BASIS: 128
BANDWIDTH: 0.082
SPTPCA VARIANCE: 220161.000 SPECTRUM TOTAL: 220161.000

FIG. 17



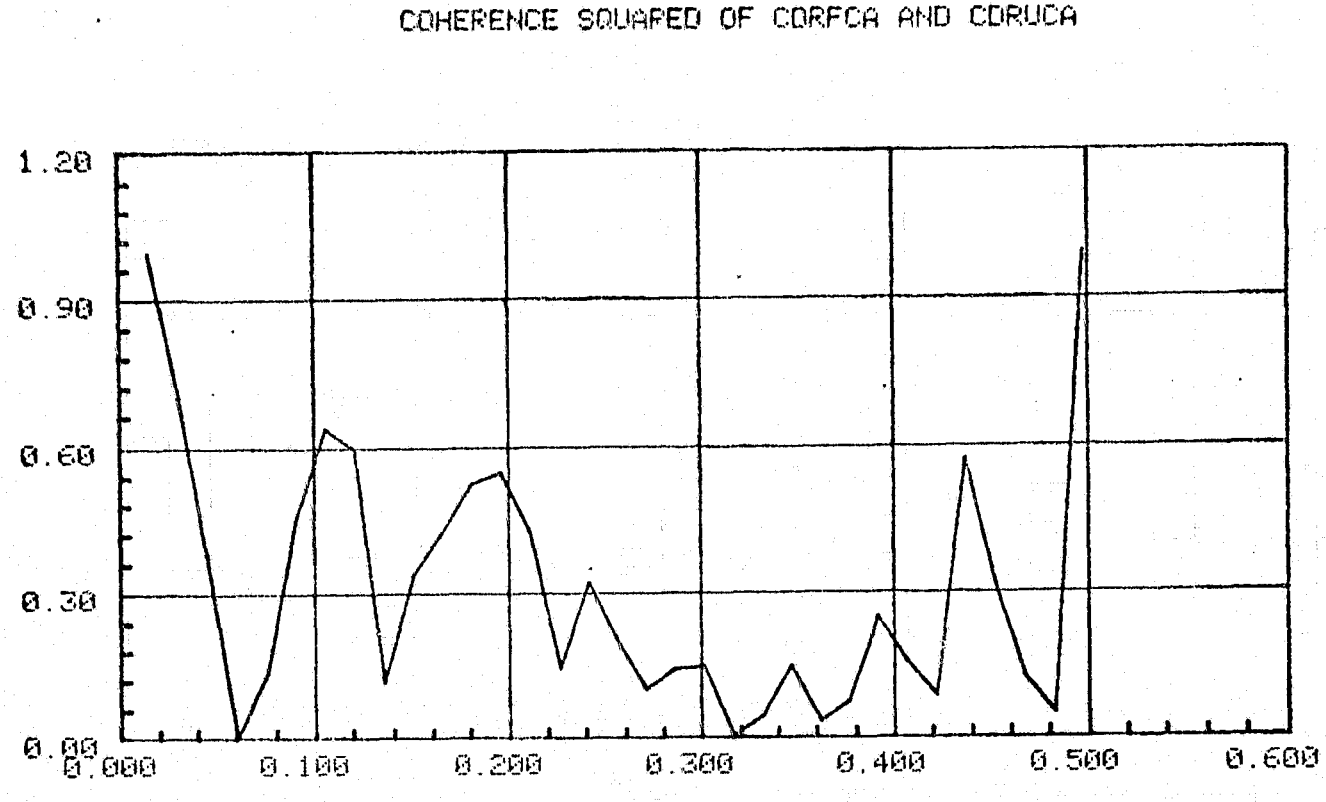

SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR  RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1  0.80
FFTPARAMETER: 84
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 84
BASIS: 84
BANDWIDTH: 0.063
NOALIGNMENT
CRITICAL COHERENCE SQUARED (AT 5 PERCENT LEVEL): 0.845
SPTFCA VARIANCE: 140055.000  SPECTRUM TOTAL: 140055.000
SPTUCA VARIANCE: 617.460  SPECTRUM TOTAL: 617.456

FIG. 18




SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 84
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 84
BASIS: 84
BANDWIDTH: 0.063
NOALIGNMENT
CRITICAL COHERENCE SQUARED (AT 5 PERCENT LEVEL): 0.845
SPTFK VARIANCE: 70805.100 SPECTRUM TOTAL: 70805.000
SPTUK VARIANCE: 506098.000 SPECTRUM TOTAL: 506094.000

FIG. 19




SUMMARY STATISTICS AND OPTIONS
TYPE OF SMOOTHING: TRIANGULAR RANGE: 4
TYPE OF PREWHITENING AND PARAMETER: TYPE1 0.80
FFTPARAMETER: 66
PREPROCESSING: DEMEAN
NUMBER OF OBSERVATIONS IN DATA SERIES: 66
BASIS: 66
BANDWIDTH: 0.080
NOALIGNMENT
CRITICAL COHERENCE SQUARED (AT 5 PERCENT LEVEL): 0.845
CORFCA VARIANCE: 232952.000 SPECTRUM TOTAL: 232951.000
CORUCA VARIANCE: 1616.770 SPECTRUM TOTAL: 1616.750

# APPENDIX II

*Estimates of Measurement Errors' Contribution to $\hat{\sigma}^2$(NET)*

Our Source Material. This appendix makes use of additional data not mentioned in the text and refers to the Measurement Document by Arnold [1]. He proposed a model for combining the means and variances of certain variables of the experiment so as to yield numbers representing the mean and variance of the net substance change per day for any astronaut. There were seven "substances"--6 minerals and nitrogen (N). Bot input and output involved a number of parameters. This appendix considers only nitrogen (N) and calcium (Ca). Measurement tables yielding quantities relevant to the model were available to us. Those we used are referenced. Their acronyms for our convenience in referencing are also included.

What We Did. We have done what we believe can be regarded as a representative calculation. We chose to particularize our calculation to the case of

Cmdr. Crippen
Day 2
Nitrogen and Calcium

only.

The claim that these calculations are representative is one we do not substantiate in detail here, and indeed going through the calculation for other substances, days, and men would be necessary to confirm this firmly. Our conclucion is a sufficiently decisive one as to the small contribution of measurement error to the total variance that we do not regard this as a crucial matter.

We have simplified Arnold's input model and changed his output model so as to allow using the data in our possession. This involved our making

reasonable estimates (described below) for a few necessary parameters based in one way or another on the data.

Our Plan in this Note. We shall state our final conclusion first, then go to particulars. The experiment, and hence the model, had three "phases"--input, urine and feces. For each phase we describe our model, how we obtain the parameter values necessary for the model, and identify any assumptions and what they are based on as they are needed.

It will turn out that where we had to make reasonable approximations, the outcome will show itself not to turn on the particular values we obtained thereby. Finally, we do the actual arithmetic to get the mean and variance of the substance--input or output for each phase.

We perform the entire sequence of calculations for each of N and Ca in turn. The following sections relate to the N findings.

## Principal Conclusions

Statements. We found the magnitudes in the measurement uncertainties associated with each phase to be in the proportions approximately of

input:feces:urine
1:10:100

Moreover, the analysis showed that almost all this ± 60% has its source in a single measurement--the concentration of substance in the urine.

Confidence in the Findings. We return to these three points at the conclusion, but these are essentially the reasons we regard our findings as trustworthy.

i) it confirms quantitatively the intuition of experts

ii) the consistency in relative and absolute orders-of magnitude of our derived-from-the-data parameters and the pre-experiment Arnold [1] values, making us feel that our estimates based on the data represented a refinement of the representative values to the particular experimental procedure used in the MBE.

iii) as an internal confirmation, we shall show moreover that any approximations we are obliged to make occur in a role influencing only negligibly the final answers.

Input

Model Variables. Conforming to Arnold's [1] (p.59) notation, we define, for a particular food the random variables:

P ≡ substance (N in this calcuation) content of portion

A ≡ substance content of standard size portion

C ≡ ratio of true weight of portion to weight of standard portion

D ≡ weight or residue when portion is judged completely consumed

We find it convenient to introduce the symbols (not appearing in Arnold [1])

H = total (rehydrated) weight of standard portion

(In Arnold's notation, $H = S + K \cdot W$, but this definition and the form in which the referenced tables are given allow us to avoid any direct mention of these variables on the right hand side.)

More imput variables ($M_1$,$M_2$) are given in Arnold (p.59), but we make an assumption that removes any necessity to consider them.

*Assumption.* All foods were completely consumed.

*Basis for Assumption.*

(1) NASA has asserted to us that in all three Skylab missions only about 10 times did it occur that a food was reported as not being completely consumed.

(2) In retrospect, given our results, this assumption cannot change the character of the conclusions from our numbers.

Relations Among Parameters. We shall find it useful (as does Arnold, p.12) that if X,Y are independent random variables, then.

(i) $E(X \cdot Y) = E(X) \cdot E(Y)$ and

(ii) $\sigma^2(X \cdot Y) = [\sigma^2(X) + E^2(X)] \cdot [\sigma^2(Y) + E^2(Y)] - E^2(X)\ E^2(Y)$

which may be rewritten as

$$= \sigma^2(X) + E^2(X)]\ (\sigma^2(Y)) + \sigma^2(X)\ E(Y)$$

Note. Proof of (ii). (ii) follows from (i), the definition of variance for any random variable Z:

$$\sigma^2(Z) = E(X^2) - E^2(Z),$$

and the fact that X,Y independent implies

$$E(X^2Y^2) - E(X^2)\,E(Y^2).$$

Remark on our presentation. Subject to our modifications, our references to Arnold [1] should be construed as for brevity's sake. We are not leaning on the prior document in any substantive way. Rather, we have reconsidered independently the most sensible way to model the input and find that it overlaps sufficiently with the previous model that will save space and be clearer if we refer to it. The way we obtain particular numerical values for some of these parameters will be entirely independent of Arnold. At the end we compare.

Explicit formulas of the model. We define as does Arnold

B = ratio of weight of portion consumed to weight of standard portion

We shall find it convenient to introduce the notation $R = \frac{D}{H}$. R is the fraction of the total portion weight that is residue (when the portion is judged completely consumed). We take, as does Arnold, H to be non-random, so that

$$E(R) = \frac{E(D)}{H}$$

We have

$$B = C - R \text{ (Arnold, p.11)} \qquad \text{and}$$

$$P = A \cdot B \text{ (Arnold, p.10)}$$

so by (i) and (ii) therefore, the expected value of the intake is

$$(1)\quad E(P) = E(A) \cdot [E(C) - E(R)]$$

and its variance is

$$(2) \qquad \sigma^2(P) = \begin{Bmatrix} [\sigma^2(A) + E^2(A)]\ \sigma^2(C) \\ + [\sigma^2(A) + E^2(A)] \cdot \sigma^2(R) \\ + \sigma^2(A) \cdot [E(C) - E(R)]^2 \end{Bmatrix} \simeq \begin{Bmatrix} E^2(A) \cdot \sigma^2(C) \\ + E^2(A) \cdot \sigma^2(R) \\ + E^2(C) \cdot \sigma^2(A) \end{Bmatrix}$$

We suppress subscripts throughout, but it must be remembered that each variable is associated with one particular food portion, i.e., (1) and (2) give the amount and variance of substance intake for each portion. To find the amount and variance of total intake for a particular day, we add those for each portion in the day's diet. Thus, the problem reduces to finding values for each food of the parameters on the right hand dise of (1) and (2). We now say how we arrived at values for each of these.

Estimating Parameters.

*Parameter:* $E(C)$

*Assumption:* $\hat{E}(C) = 1$

*Basis:* (1) The standard weights were asserted to us to have been "unbiased," i.e., roughly as likely to be about as much above or below the true portion weights for any portion. (2) Same as (2) of assumption on page 3 of this Appendix.

*Parameter:* $\sigma^2(C)$

*Assumption:* $\hat{\sigma}(C) = 0.01$ for all foods.

*Basis:* We were told in personal conversation that NASA arbitrarily adopted the standard of keeping each food's actual weight of standard portion to within ± 2 to 2.5% of the specified standard portion weight. (Examining preliminary data tells us there may have been a few exceptions to this.) But taking 2% = 0.02 to be 2 standard deviations ($2\sigma$), we are led to the assumed estimate below.

*Parameter:* $E(A)$. $\hat{E}(A)$ may be read directly from the 9 sample table under the column headed MEAN, in the row for nitrogen--one value for each food.

*Parameter:* $\sigma^2(A)$. $\hat{\sigma}^2(A)$ comes from the same row of the same table, as the square of the item in the column STD DEV. We have found it most natural to obtain directly from the data the mean and variance of the parameter R, as follows:

*Parameters:* $E(R)$, $\sigma^2(R)$. To avoid unnecessary calculation and to deal with the fact that some foods occur so seldom in the diet that individual values of $E(R)$ and $\sigma^2(R)$ could not be based upon a satisfactorily large sample, we make the following assumption.

*Assumption:* Any two foods contained in the same container-type (can) share the same values of $E(R)$ and $\sigma^2(R)$.

*Basis:* (1) An examination of Table 9 in Arnold would show that the table was prepared under the same assumption. This is in accord with reasonable common sense. (2) Retrospective order-of-magnitude examination of the model equation for the input variance ($\sigma^2(P)$) will show the variable $E(R)$ to be the least important variable affecting the final variance-of-net-change of any quantity in the model. Its influence on the expected value of the net change will be shown to be inconsequential as well. (3) Examination of actual data values comparing observed $E(R)$, $\sigma^2(R)$'s for different foods in the same type container from the residue table showed that this assumption held to high accuracy.

How does this assumption work out in practice? Examining Table 9 in Arnold shows the counts for number of foods for each can type (Chart 1).

Chart 1. Number of Foods in Diet in Each Container Type

| TYPE | NUMBER OF FOODS |
|---|---|
| I | 2 |
| II | 2 |
| III | 1 |
| IV | 5 |
| V | 16 |
| VI | 4 |
| VII | 26 |
| VIII | 11 |
| IX | 1 |
| X | 1 |
| XI | 0 |
| XII | 1 |

*Assumption:* For foods in container of Types I, IV, VI, IX and X.

$$E(R) = 0, \quad \sigma^2(R) = 0$$

*Basis:* Examining the residue table shows these foods, being generally "hard" non-viscous substances, to have vanishingly small residues, which we attribute to the measurement procedure rather than to food really leftover. Examining Chart 1 shows that most foods were in containers either of Types V, VII or VIII. Thus, there are ample observations in the residue table from which to obtain $\hat{E}(R)$ and $\hat{\sigma}^2(R)$ directly from the data for each of these three can types.

Our method was to pick 20 foods (informally) at random from Crippen's in-chamber days, allowing for the possibility that $E(R)$, $\sigma^2(R)$ might be different for different astronauts. Columns headed "residue" and "total" are values of D and H, respectively. We divided the D for each of our sample portions by its associated H to find $R_i$, say. We then set

$$E(R) = \frac{1}{20} \sum_{i=1}^{20} R_i \equiv \bar{R}$$

and

$$\sigma^2(R) = \frac{1}{20} \sum_{i=1}^{20} (R_i - \bar{R})^2$$

We thought 20 foods would be a large enough sample to obtain accurate enough values for $\hat{E}(R)$, $\hat{\sigma}^2(R)$. Chart 2 shows the calculation for can type VII with the decimal point moved over for convenience in the last column.

Note. Here, as in all the charts in this Appendix, we include an exaggerated number of significant figures in our numbers. This is because we regard these quantities as intermediates toward the computation of E and $\sigma^2$ for NET, and it is standard numerical-analytic common sense to round off at the end, not before.

For other less common can types, observations were harder to come by. This is not a flaw to be worried over, however, if we remember that it is exactly these foods that occur less often in the diet so their parameter values entered less often. Can types III and XII occurred so rarely that it was not feasible to keep their parameters astronaut-specific.

Chart 2. Crippen (in-chamber values) Can Type VII

| D | H | $\frac{D}{H}$ | $\left(\frac{D}{H} - \frac{E(D)}{H}\right)^2 \cdot 10^8$ |
|---|---|---|---|
| 0.79 | 68.4 | 0.0115 | 8 |
| 1.07 | 124.5 | 0.0086 | 1011 |
| 0.22 | 75.0 | 0.0029 | 7885 |
| 0.42 | 124.5 | 0.0034 | 7022 |
| 0.84 | 171.8 | 0.0049 | 4733 |
| 6.28 | 110.5 | 0.0568 | 202680 |
| 2.18 | 172.5 | 0.0126 | 67 |
| 0.00 | 167.6 | 0.0000 | 13924 |
| 3.08 | 199.0 | 0.0155 | 1384 |
| 2.35 | 156.0 | 0.0151 | 1102 |
| 1.38 | 68.4 | 0.0202 | 7089 |
| 1.17 | 124.5 | 0.0094 | 576 |
| 0.02 | 75.0 | 0.0003 | 13225 |
| 1.50 | 196.5 | 0.0076 | 1764 |
| 2.09 | 172.5 | 0.0121 | 9 |
| 1.37 | 110.5 | 0.0124 | 36 |
| 1.67 | 124.5 | 0.0134 | 256 |
| 1.05 | 199.2 | 0.0053 | 4225 |
| 0.91 | 183.0 | 0.0050 | 4624 |
| | | 0.2356 | 276244 |

TOTALS: $\frac{\hat{E}(D)}{H} = 0.01178$ $\frac{\hat{\sigma}^2(D)}{H^2} = 13812 \cdot 10^{-8}$

Chart 3 shows the values of $\hat{E}(R)$ and $\hat{\sigma}^2(R)$ computed by this scheme, and the number of observations in the sample for each can type.

This completes our specification of the input calculation.

Chart 3. n = Number of Foods Used to Compute Statistics

| CAN TYPE | $10^2 \cdot E(R)$ | $(1-E(R))$ | $(1-E(R))^2$ | $10^4 \cdot (\sigma^2(R))$ | n |
|---|---|---|---|---|---|
| II | 2.090 | 0.9791 | 0.9586 | 5.79 | 8 |
| III | 0.424 | 0.9958 | 0.9916 | 0.0321 | 5 |
| V | 0.772 | 0.9923 | 0.9847 | 0.1571 | 20 |
| VII | 1.178 | 0.9882 | 0.9765 | 1.381 | 20 |
| VIII | 2.269 | 0.9773 | 0.9551 | 1.777 | 20 |
| XII | 2.380 | 0.9762 | 0.9530 | 5.08 | 9 |

Can types I, IV, VI, IX, X:

$$E(R) = \sigma^2(R) \equiv 0.$$

Model Calcuations. Having specified all parameters and described how we arrived at a numerical value of each, we now carry out explicitly the arithmetic dictated by the model calcuations. Recall again that we are focusing on day 2 of Cmd. Crippen (See Arnold, p.53 for this diet menu on that day).

Chart 4 exhibits the contributions of each portion on day 2 to the total input amount ($E(P)$) and variance ($\sigma^2(P)$) both according to "term of the sum" in the $\sigma^2(P)$ equation and "food." Chart 4 and the arithmetic summing totals for each column shows several findings.

Findings.

1. Referring to Chart 3 for the values of $\hat{E}(R)$ and its square, we see from the closeness of the values for different can types that a uniform value for all types would have sufficed in the $E(P)$ column.

2. A small fraction of the foods contribute almost all the N, and these same foods contribute almost all the uncertainty in N.

Chart 4. Contributions of Each Food in Crippen's Day 2 Diet to E(P) [eqn.(1), p.4], and to Each of the Three Terms [eqn.(2), p.5] for $\sigma^2(P)$, the Latter Shown $\times 10^6$ to Avoid a Clutter of Zeros

| Food No. | E(P) | $10^6 \times [\sigma^2(A)+E^2(A)]\cdot\sigma^2(C)$ | $10^6 \times [\sigma^2(A)+E^2(A)]\frac{\sigma^2 D}{H^2}$ | $10^6 \times \sigma^2(A)\cdot[E(C) - \frac{E(D)}{H}]^2$ |
|---|---|---|---|---|
| 38 | 0 | 0 | 0 | 0 |
| 16 | 2.2175 | 504.6 | 696.85 | 9570.68 |
| 13 | 2.5249 | 653.4 | 902.35 | 5492.81 |
| 8 | 0.0801 | 0.7 | 1.24 | 0.96 |
| 8 | 0.0801 | 0.7 | 1.24 | 0.96 |
| 62 | 0.0723 | 0.5 | 0.89 | 0 |
| 62 | 0.0723 | 0.5 | 0.89 | 0 |
| 71 | 4.9074 | 2466.1 | 3405.68 | 47.85 |
| 65 | 0.5583 | 31.9 | 44.05 | 249.98 |
| 47 | 0 | 0 | 0 | 0 |
| 19 | 0.5580 | 31.7 | 4.98 | 15.75 |
| 42 | 0 | 0 | 0 | 0 |
| 42 | 0 | 0 | 0 | 0 |
| 62 | 0.0723 | 0.5 | 0.89 | 0 |
| 39 | 3.7400 | 1420.9 | 223.22 | 3908.27 |
| 29 | 0.5089 | 26.5 | 36.60 | 24.41 |
| 69 | 0.3666 | 13.8 | 19.06 | 316.39 |
| 8 | 0.0801 | 0.7 | 1.24 | 0.96 |
| 66 | 0 | 0 | 0 | 0 |
| 66 | 0 | 0 | 0 | 0 |
| 62 | 0.0723 | 0.5 | 0.89 | 0 |
| 62 | 0.0723 | 0.5 | 0.89 | 0 |
| 62 | 0.0723 | 0.5 | 0.89 | 0 |
| 4 | 0 | 0 | 0 | 0 |
| 42 | 0 | 0 | 0 | 0 |
| 42 | 0 | 0 | 0 | 0 |
| 60 | 0 | 0 | 0 | 0 |
| TOTALS | 16.056 | 5744.5 | 5425.56 | 35008.01 |

Thus,

$$E(P) = 16.056$$

$$\sigma^2(P) = (5.7445 + 5.42556 + 25.00801 \times 10^{-3} = 0.046178 \text{ g}$$

Implication of (2). Some preliminary experiments could have been done to see which foods contribute negligibly to each substance's total, thus reducing the number of measurements that have had to have been carried along through the experiments themselves, for N at least.

Output--Feces

Model Variables. We define:

$N_f$ = g of nitrogen/fecal sample

$p$ = percent of sample that is N

$f$ = g of feces in sample

Our model uses somewhat differently defined variables than in Arnold although it is the same in concept, namely that

(*) total substance = $\begin{pmatrix}\text{fraction of sample} \\ \text{that is substance}\end{pmatrix}$ · (size of sample)

Thus, we have

$$N_f = p \cdot f$$

Relations Among Parameters. We need not involve equation (i) to obtain $E(N_f)$ because this quantity may be read directly from the daily means table in the N column for each day. Using equation (ii) implies

$$(5) \qquad \sigma^2(N_f) = \sigma^2(p)\ \sigma^2(f) + \sigma^2(p)\ E^2(f) + E^2(p)\ \sigma^2(f)$$

Estimating Parameters. We now tell how we obtained estimates of each quantity on the right hand side of (5)

*Parameter:* $E(p)$. $\hat{E}(p)$ was obtained directly from the p-% duplicates table simply by averaging both members of every pair for many pairs spread over all three phases of SMEAT (pre-, in-, and post-chamber). Averaging a sample of about one-third the total number of observations in the table, we find $\hat{E}(p)$ - 4.204058.

*Parameter:* $\sigma^2(p)$. We may estimate this because values in the p-% duplicate table are in duplicate.

If $p_{i_1}$, $p_{i_2}$ are the random variables representing the duplicate determinations of p on the ith sample, or the actual sample values depending on our context, and if

$$\overline{p} = \frac{p_{i_1} + p_{i_2}}{2}$$

then

$$\sigma^2(p_{i_1}) = \sigma^2(p_{i_2}) = E(p_{i_1}) - E(p_{i_1})_2$$

Because $E(p_{i_1}) = \overline{p}_{i_1}$, we may take as the sample estimate

$$\hat{\sigma}^2(p_i) = (p_i) - \frac{p_{i_1} + p_{i_2}}{2})_2$$

which is equal to

(6) $$\frac{(p_{i_1} + p_{i_2})^2}{4}$$

For a set of samples with sample numbers {i} over which $\sigma^2(p)$ is constant, an appropriate value for $\hat{\sigma}^2(p)$ would therefore be the average of quantities (6) over all indices in the set {i}.

Thus, if we were guaranteed that $\sigma^2(p)$ would not change over our entire range of samples, we would take the differences between members of every duplicate pair, square them and divide by 4, and the average of all these would be our estimate of $\sigma^2(p)$. Are we justified in assuming $\sigma^2(p)$ constant over the entire range of sample numbers in the p-% duplicates table? Examining the data shows we definitely are not. Taking subsamples from three locations in the table reveals the sample variance estimates obtained by the above scheme in each section as in Chart 5.

Chart 5

| Sample Nos. | Number in Sample | $10^4 \times \sigma^2(p)$ |
|---|---|---|
| 42-57 | 13 | 0.00166 |
| 81-96 | 14 | 0.00165 |
| 222-233 | 10 | 0.01580 |

Thus, for a reason we do not know nor find a need for this analysis, the variability in the highest sample numbers is an order-of-magnitude greater than that in the rest of the measurements.

*Assumption.* A representative value for $\sigma^2(p)$ in the entire sample may be obtained by taking a weighted average of thw to estimates from Chart 5, namely $10^{-4} \times 0.00166$ and $10^{-4} \times 0.01580$, weighted 11:5, respectively.

*Basis for Assumption.* SMEAT's timetable was

| Phase | No. of 6-day Periods/Phase |
|---|---|
| pre-chamber | 4 |
| in-chamber | 9 |
| post-chamber | 3 |

The high sample numbers, we suspect, were obtained during the post-chamber phase, and because the pre-chamber periods tended to be a troublesome time in SMEAT, we attach the larger variance-figure to half (arbitrarily) the pre-chamber period as well as to the post-chamber period, making the appropriate weights on the smaller and larger variance

$$\frac{16-t}{16} \quad \text{and} \quad \frac{t}{16} \ ,$$

where $t = 3 + \frac{1}{2}(4) = 5$.

Thus, we take as our estimate of $\sigma^2(p)$ for SMEAT the value

$$\sigma^2(p) = \frac{(11)(0.00166) + (5)(0.01580)}{16} \times 10^{-4}$$

$$= 0.006079 \times 10^{-4}$$

C.2

*Parameter.* E(f). Here we had the problem that we were not provided with the correspondence in labeling between the two tables, daily means and p-% duplicates, and also were not given any values at all for f. We do not view this as a serious drawback, however, because we are doing a representative calculation here and what is therefore important are representative values which we obtain by averaging or "smoothing." Because of the irregular and discrete quality of fecal sampling, some smoothing would almost inevitably be desirable in any statistical evaluation of these results even in the presence of full data on the variable f. Thus, we regard the following averaging to yield a value for E(f) as quite satisfactory and trustworthy for this evaluation.

Although we are not given values on f itself, we are given values on a fecal variable besides $N_f$ and p, namely

K = amount N per fecal determination (g)

Our plan is to obtain from the daily means table a suitable average value of K, say $\hat{K}$, and then take

$$\hat{E}(f) = \frac{\hat{K}}{\hat{E}(p)} \tag{11}$$

Multiplying both sides by $\hat{E}(p)$, this simply says

$$\left(\begin{array}{c}\text{representative value of}\\ \text{expected size of fecal}\\ \text{sample}\end{array}\right) \cdot \left(\begin{array}{c}\text{\% of fecal sample}\\ \text{that is N}\end{array}\right) = \left(\begin{array}{c}\text{average amount N}\\ \text{in a fecal sample}\end{array}\right) \tag{7}$$

To obtain $\hat{K}$, we compute initially three separate values for the three experimental phases because evacuation habits can be presumed to differ in the phases. If $K_i$ is the value of K on the $i^{th}$ fecal determination, then

$$\hat{K}_{pre} = \frac{\sum_{i=1}^{15} K_i}{15} = \frac{24.470}{15} = 1.6313$$

$$\hat{K}_{in} = \frac{\sum_{i=16}^{50} K_i}{35} = \frac{48.528}{35} = 1.3865$$

$$\hat{K}_{post} = \frac{\sum_{i=51}^{60} K_i}{10} = \frac{22.960}{10} = 2.2960$$

thus showing that stratifying by phases was indeed called for. We take $K$ as the weighted average of these three phase-specific estimates, with weights proportional to the fraction of SMEAT in each phase:

$$\hat{K} = \frac{4\,\hat{K}_{pre} + 9\,\hat{K}_{in} + 3\,\hat{K}_{post}}{16}$$

yielding

$$\hat{K} = 1.6182 \text{ g}$$

whereby from (11) we find

(12a) $$\hat{E}(f) = \frac{1.6182}{4.204058 \times 10^{-2}} = 38.4914$$

and

(12b) $$\hat{E}^2(f) = 1481.59$$

Because we want to do the calculation for a <u>one-day</u> period and because feces appear irregularly, we need a value for mean g N from feces <u>per day</u>. This may be obtained by smoothing $\hat{K}$ to find this number as if there were a fecal sample of the same size every day fo SMEAT. We have $\hat{E}(N_f)$ equal to

$$\frac{\text{mean g N}}{\text{day}} = \frac{\text{mean g N}}{\text{determination}} \cdot \frac{\text{number of determinations}}{\text{number of days}}$$

$$= \hat{K} \cdot \frac{60}{(6)(16)}$$

because there were 60 fecal samples in SMEAT (i goes up to 60 in computing the $\hat{K}$'s). There were 16 6-day periods in SMEAT.

Model Calculations. Computation yields

$$\hat{E}(N_f) = 1.0114 \text{ g/day}$$

By (5), we have

$$10^4 \times \hat{\sigma}^2(N_f) = (0.006079)\,\sigma^2(f) + (0.006079)(1481.59) \cdot \left(\frac{60}{96}\right)^2 + (17.674104)\,\sigma^2(f)$$

or

$$(9) \quad \hat{\sigma}^2(N_f) = [(17.68017)\,\sigma^2(f) + 3.51797] \times 10^{-4}.$$

*Assumption:*

$$(8) \quad \hat{\sigma}^2(f) = 0.01\,\hat{E}(f).$$

*Basis for Assumption:*

(1) If, as might seem reasonable, f is measured to within ± 2% and we take this to be "2σ," then equation (8) holds. If the figure 2% were not itself so arbitrary, we should multiply the $\hat{\sigma}^2(f)$ in (8) by $\left(\frac{60}{6(16)}\right)$ to obtain an "equivalent daily" standard deviation, but nothing crucial hinges on this admustment as we

(2) shall see retrospectively: the fecal measurments are the most accurate part of the experiment and the exact number here will influence the final answer only negligibly. Equation (8) implies that

$$\hat{\sigma}^2(f) = 10^{-4}\,E^2(f) = [0.14816\,\left(\frac{60}{96}\right)^2] = 0.05788$$

so the first term in brackets in (9) becomes with this assumption

$$(17.68017)(0.05788) = 1.02333$$

which is considerably smaller (although not negligible) with respect to 3.51797 even still in this "intra-fecal" analysis.

We find with this assumption

$$\sigma^2(N_f) = 4.5413 \times 10^{-4}\ g^2$$

Output--Urine

Model Variables. We define:

$N_u$ = g N from urine per day

c = concentration of N in urine (mg/100 cc)

V = daily urine volume (ml)

Relations Among Parameters. Again, (*) expresses the idea relating the variables:

$$N_u = (10^{-5}c) \cdot V$$

The factor comes by remembering that 1 ml = 1 cc

From (i) and (ii),

$$E(N_u) = E(c) \cdot E(V) \cdot 10^{-5}$$

$$\sigma^2(N_u) = \{\sigma^2(c)\ \sigma^2(V) + \sigma^2(c)\ E^2(V) + E^2(c)\ \sigma^2(V)\} \times 10^{-10}$$

Estimating Parameters.

*Parameter:* E(c). The c-duplicates table gives duplicate determinations and their means. Examining the data shows no strong trend over time, so we may be satisfied with averaging many c-values spread over the whole of the experiment to obtain $\hat{E}(c)$. Doing this yields

$$\hat{E}(c) = 549.06\ \frac{mg}{100\ cc}$$

*Parameter:* $\sigma^2(c)$. We compute the quantity

(13) $$[(\text{larger value}) - (\text{mean of pair})]^2$$

for 20 pairs of Crippen's in-chamber c-values. In-chamber was chosen because, looking at the experiment as a whole, this appears to have been the most stable phase. Thus, the figure we obtain may be slightly on the small side. We obtained these 20 values of (13) (A number in parentheses counts how many times the number preceding occurred.) Ordered by size, they are: 1(3); 10, 25, 36, 49, 64, (2); 81, 100, 144 (2); 225 (2); 256, 361, 784 (2); 1024. Their average is

$$\sigma^2(c) = 218.9\ (\frac{mg}{100\ cc})^2$$

*Parameter:* $E(V)$. We obtain a value for $\hat{E}(V)$ in the same way as for $\hat{E}(c)$--by averaging in-chamber V values.

We obtain

$$\hat{E}(V) = 2499 \text{ ml}$$

*Parameter:* $\sigma^2(V)$

*Assumption:*

$$\hat{\sigma}^2(V) = 10 \text{ ml}^2$$

*Basis:* We are not given any direct information about the measurement accuracy of V, but we can surmise from examining (a) the significant figures in their sample values and (b) the units-digits in these numbers that they have been measured to within about ± 5-7 ml. Taking the midpoint of this interval, 6, to be "2σ," then "σ" = 3 and "$\sigma^2$" = 9, which we found off to 10. [See section on Findings below.]

Model Calculations. Thus, we may compute

$$\hat{E}(N_u) = (549.06)(2499)\ 10^{-5} = 13.721 \text{ g}$$

and

$$\hat{\sigma}^2(N_u) = \{\hat{\sigma}^2(c)\ \hat{\sigma}^2(V) + \hat{\sigma}^2(c)\ \hat{E}^2(V) + \hat{E}^2(c)\ \hat{\sigma}^2(V)\}\ 10^{-10}$$

$$= \{(218.9)(10) + (218.9)(2499)^2 + (549.06)^2(10)\}\ 10^{-10}$$

$$= \{0.002189 + 1367.8 + 3.0147\} \times 10^{-4}$$

$$= 1370.8 \times 10^{-4} \quad \text{or} \quad \hat{\sigma}^2(N_u) = 0.1371 \text{ g}^2$$

Findings. Essentially the entire variance in urine N comes from the term

$$\hat{\sigma}^2(c)\ \hat{E}^2(V)$$

Thus, still within this "intra-urine" analysis, we may conclude that the exact value chosen for $\sigma^2(V)$ will influence negligibly the total uncertainty in urine output. We claimed this previously.

Implication. So far as establishing measurement error goes, comparatively little attention in the experimental procedure need be paid to especially accurate measurement of V and essentially all attention should be focused on making $\sigma^2(c)$ as small as possible.

Net Intake

Model Variables. Here we combine the "intra" analyses to find the mean and measurement variance of the net intake, say NET. Estimating this quantity is the goal of the MBE. We have

$$NET = P - N_f - N_u, \qquad \text{so}$$

Relations Among Parameters.

(15) $$E(NET) = E(P) - E(N_f) - E(N_u)$$

(16) $$\sigma^2(NET) = \sigma^2(P) + \sigma^2(N_f) + \sigma^2(N_u)$$

Estimating Parameters and Model Calculations. Thus, using our "intra" results

(17) $$E(NET) = 16.056 - 1.0114 - 13.721 = 1.3236 \text{ g}$$

(18) $$\sigma^2(NET) = 0.046178 + 0.00045413 + 0.13708$$

$$\sigma^2(NET) = 0.18371 \quad \sigma(NET) = 0.42862$$

(19) $$E(NET) \pm 2\sigma(NET) = + 1.3236 \pm 0.8572$$

is thus a one-day MBE confidence interval associated with measurement error only. Its scope of inference is the restricted one associated with "measurement error only" as described in the text under the sections headed "The Restricted Inference" and "Status of the Arnold Document [1]." The confidence interval given above in (19) is not quite overlapping zero, but close enough that confidence even in the direction of net change (i.e., did he gain or lose the substance?) must be very weak. Confidence in the numerical result itself is totally unwarranted. $2\sigma(NET)$ is 65% of $E(NET)$.

Now notice from where this uncertainty comes. Chart 6 shows the percentage of the total variance contributed by the variances from each phase of the experiment.

Chart 6

PHASE

| Input | Feces | Urine |
|---|---|---|
| 25.0 | 0.6 | 74.3 |

Now recalling the calculation of $\hat{\sigma}^2(N_u)$, we found that $\hat{\sigma}^2(N_u)$ is due almost entirely to the term involving $\hat{\sigma}^2(c)$. Our conclusion, therefore, is that the measurement of substance concentration in urine is contributing three-quarters of the measurement error.

Magnitudes of the uncertainty in the fecal, input and urine phases of the experiment go like

$$\frac{1}{50} : 1 : 3$$

The calculation of E(NET) shows that far more nitrogen is coming from the urine than from the feces. Although in principle not relevant to the variance calculation because the model assumes homogeneity of all variances (i.e., independent of the size of the variable being measured), this is for mathematical simplicity, and our assuming it does not make it true; in fact such dependence is quite common.

In the presence of any such dependence, the fact that most output nitrogen comes in the urine further underscores the importance of reducing $\sigma^2(c)$ in any MBE where restricted inferences are important, i.e., in any experiment whose particular outcome is important (again as distinguished from inferences with the wider scope as described in the text).

Confidence in the Findings.

1) Finding (2) gives quantitative support to the intuition of experts in the field of biomedical experiments that this measurement is an uncertain one. The present calculation shows that this one measurement produces a final uncertainty that is an order-of-magnitude greater than that generated by the entire input phase.

2) Our approximations have been sensible and conservative, and have been shown even "intra"-phase not to influence the answer much in any case; this is highly strengthened when we view how negligible their effects are on the net answers.

3) Our parameter estimates were computed on the basis of the SMEAT data themselves and are thus "tailored to this particular experiment." This view, that our "hatted" quantities are refinements of estimates of these quantitites appearing in laboratory manuals, or otherwise able to be anticipated approximately before the experiment, is substantiated by Chart 7, which exhibits the consistency in orders-of-magnitude of our estimates with those in Arnold, these latter representing the best available estimates of these parameters before SMEAT.

Chart 7. Output Parameter Estimates Calculated by this Paper from Pre-Skylab Data, compared with Arnold's values

| Parameter | SMEAT Value, $g^2$ (Arnold) | Our "Hatted" Estimate |
|---|---|---|
| $\sigma^2(p)$ | $8.19925 \times 10^{-7}$ (p.48) | $6.079 \times 10^{-7}$ |
| $\sigma^2(f)$ | $2.5 \times 10^{-5}$ (p.48) | $14816 \times 10^{-5}$ |
| $\sigma^2(c)$ | $0.7050 \times 10^{-8}$ (p.49) | $2.189 \times 10^{-8}$ |
| E(f) | 20.74 (p.57) | 38.49 |
| E(c) | 9.00 mg/ml (p.58) | 5.49 mg/ml |
| E(V) | 2311 ml (p.58) | 2499.7 ml |
| $\sigma^2(V)$ | Not given | 10 |
| E(p) | 0.050 g/g (p.57) | 0.04204 g/g |

Discussion of Chart 7. (a) The two parameters with sharply discrepant values are $\sigma^2(f)$ and $\sigma^2(c)$. We arrived at a value of $\sigma^2(f)$ very arbitrarily and Arnold is therefore telling us that we were highly overly pessimistic. But because the fecal phase even with the pessimistic assumption is the most accurate in the experiment, we do not dwell upon this discrepancy.

(b-i) The other discrepancy involves $\sigma^2(c)$, the most important parameter in the calculation. We would be more suspicious of our computed value of $\sigma^2(c)$ as being somehow unrealistically large, were it not for the fact

that we explicitly acknowledged a possible bias in favor of the well-controlled period of SMEAT, the in-chamber period.

Thus, we conclude that our finding shows that the c-measurements not to have been as accurate as it was believed prior to SMEAT that they could be. (b-ii) The character of our findings remain unchanged with the Arnold $\sigma^2(c)$ value instead of our $\hat{\sigma}^2(c)$, although now the terms in which one would describe the finding become a little more moderate.

Overview of Nitrogen Calculations

The present remarks give a quantitative measure to the relative attention that should be paid in the context of measurement error and restricted inference (see text) to different phases of an MBE, and they confirm the qualitative feeling of practitioners. They also answer a specific question about nitrogen intake.

The Corresponding Calculations for Calcium

Unlike nitrogen, the primary excretory source of the substance Ca is the feces. Because the urine mode made the largest contribution to total measurement variance for N, it is of interest to see to what extent this N result is sustained in this different Ca situation.

We begin with Table SAME which exhibits the parameters that are "substance-dependent" and therefore whose values remain the same in principle for the Ca calculation as for the N calculation.

Table SAME

Model Parameter Estimates with the Same Value in N and Ca Calculations

| Parameter | Estimate |
|---|---|
| $E(C_i)$ | 1 |
| $\sigma^2(C_i)$ | $1 \times 10^{-4}$ |
| $E(R_i)$ | As given for each can type on p. (Progress Report) |
| $\sigma^2(R_i)$ | Same as above |
| $E(V)$ | 2499.7 ml |
| $\sigma^2(V)$ | 10 $ml^2$ |
| $E(f)$, $\sigma^2(f)$ | (approximately, see text) |

The following sections summarize our computations of the remaining parameter-estimates in an order paralleling that of the N calculation.

Table $P_{Ca}$ exhibits for Ca the contribution from each food to $\hat{E}(P)$ and to each term in the sum (2) representing $\hat{\sigma}^2(P)$. Again a small fraction of the foods contribute to $\hat{E}(P)$, and these same foods also contribute most of $\hat{\sigma}^2(P)$. Thus, the implication that the number of measurements taken in the MBE could have been reduced substantially applies for Ca as well as for N .

*Parameter:* $E(P_{Ca})$. To obtain $\hat{E}(p)$ for Ca, we used the p-% duplicates table for N as follows. The average of

$$\frac{g\ Ca}{g\ N} = r, \text{ say,}$$

over many samples spread over all phases is:

$$\bar{r} = 0.4080597.$$

Multiplying this by $\hat{E}(p)$ for N yields our estimate of $\hat{E}(p)$ for Ca:

$$\hat{E}(p_{Ca}) = (0.4080597)(4.204058) = 1.715507$$

*Parameter:* $\sigma^2(p_{Ca})$. For $\hat{\sigma}^2(p_{Ca})$, we used

$$\begin{aligned}\hat{\sigma}^2(p_{Ca_{data}}) &= \hat{\sigma}^2(\bar{r} \cdot p_N) \\ &= \bar{r}^2\, \hat{\sigma}^2(p_N) \\ &= (0.4070597)^2 \cdot (6.079 \times 10^{-7}) \\ &= 0.10122308 \times 10^{-6}\end{aligned}$$

Because of the somewhat ad hoc character of the above choices, we shall also "carry along" the value of $\hat{\sigma}^2(p_{Ca})$ offered in Arnold:

$$\hat{\sigma}^2(p_{Ca_{Arnold}}) = 0.000321 \times 10^{-6}.$$

Notice that they are indeed discrepant. At the end we see how big a difference this makes in the NET variance.

Table $P_{Ca}$

| Food No. | $\hat{E}(P)$ | $10^2 \times [\hat{\sigma}^2(A)+\hat{E}^2(A)]\ \hat{\sigma}^2(C)$ | $10^2 \times [\hat{\sigma}^2(A)+\hat{E}^2(A)]\ \hat{\sigma}^2(R)$ | $10^2 \times \hat{\sigma}^2(A)[1-\hat{E}(R)]^2$ |
|---|---|---|---|---|
| 38 | 22.4954 | 5.1907 | 0.8155 | 509.19 |
| 16 | 146.2091 | 219.1243 | 304.8019 | 2122.33 |
| 13 | 10.2032 | 1.0799 | 1.5021 | 135.53 |
| 8 | 69.3433 | 50.5550 | 89.8362 | 2005.23 |
| 8 | 69.3433 | 50.5550 | 89.8362 | 2005.23 |
| 62 | 4.2601 | 0.1906 | 0.3387 | 5.92 |
| 62 | 4.2601 | 0.1906 | 0.3387 | 5,92 |
| 71 | 23.8907 | 5.8496 | 8.1368 | 45.31 |
| 65 | 3.6257 | 0.1370 | 0.1906 | 19.33 |
| 47 | 10.7139 | 1.1713 | 0.1840 | 52.58 |
| 19 | 29.0526 | 8.5737 | 1.3469 | 16.44 |
| 42 | 3.0062 | 0.0957 | 0.1701 | 6.30 |
| 42 | 3.0062 | 0.0957 | 0.1701 | 6.30 |
| 62 | 4.2601 | 0.1906 | 0.3387 | 5.92 |
| 39 | 271.3494 | 748.0230 | 117.5144 | 2432.31 |
| 29 | 69.9626 | 50.1545 | 69.7649 | 297.34 |
| 69 | 30.3160 | 9.4269 | 13.1128 | 126.26 |
| 8 | 69.3433 | 50.5550 | 89.8362 | 2005.23 |
| 66 | 0.6245 | 0.0040 | 0.0071 | 0.29 |
| 66 | 0.6245 | 0.0040 | 0.0071 | 0.29 |
| 62 | 4.2601 | 0.1906 | 0.3387 | 5.92 |
| 62 | 4.2601 | 0.1906 | 0.3387 | 5.92 |
| 62 | 4.2601 | 0.1906 | 0.3387 | 5.92 |
| 4 | 10.6751 | 1.2060 | 6.9827 | 165.26 |
| 42 | 3.0062 | 0.0957 | 0.1701 | 6.30 |
| 42 | 3.0062 | 0.0957 | 0.1701 | 6.30 |
| 60 | 81.8762 | 70.4723 | 125.2293 | 2715.35 |
| $\hat{E}(P)$ = | 957.2342 mg | 1273.6080 | 921.8173 | 14714.22 |

So

$10^2 \times \hat{\sigma}^2(p) = 16909.65\ mg^2$

$\hat{\sigma}^2(P) = 169.0965\ mg^2$

$\sigma(P) = 13.003711$ mg

*Parameter:* $\hat{E}(f)$. Although $\hat{E}(f)$ should not in principle change (see Table SAME), because of some missing values for N which were present in the Ca fecal table for CDR, we may recompute $\hat{E}(f)$ with a few additional samples to find that its value does change slightly from (12a); by (11) we have

$$E(f) = \frac{712.405 \times 10^{-3} g}{1.715507 \times 10^{-2}}$$

$$= 41.527 \text{ g}$$

*Parameter:* $\hat{K}$. Again, the few missing N values led to a slightly larger sample for the $\hat{K}$ calculation for Ca: parallel to that for N, we find Table K.

Table K

| Phase | No. of Observations | $\hat{K}$ (by phase) in mg |
|---|---|---|
| Pre-chamber | 16 | 623.35 |
| In-chamber | 38 | 710.28 |
| Post-chamber | 11 | 837.52 |

From Table K, we compute

$$K = \frac{4(623.35) + 9(710.28) + 3(837.52)}{16} = 712.405 \text{ mg}$$

*Parameter:* $E(Ca_f)$. Using $\hat{K}$, we find for $\hat{E}(Ca_f)$ corresponding to the calculation for $\hat{E}(N_f)$:

$$\hat{E}(Ca_f) = (712.405)\left(\frac{65}{6(16)}\right) = 482.358 \text{ mg}$$

*Parameter:* $\sigma^2(f)$. Assuming as we did for N that $\sigma(f) = .01\ E(f)$, we find

$$\hat{\sigma}^2(f_{data}) = 10^{-4} \times E^2(f) = 10^{-4} \times (41.527)^2 = .17245 \text{ g}^2$$

We may compare this value with the Arnold value

$$\hat{\sigma}^2(f_{Arnold}) = 2.5 \times 10^{-5} g^2$$

to find that, as for N, the former $\hat{\sigma}^2(f)_{data}$ is again probably an overly pessimistic estimate of $\sigma^2(f)$.

*Parameter:* $\sigma^2(Ca_f)$. Combining the foregoing parameter estimates according to

$$\hat{\sigma}^2(Ca_f) = \hat{\sigma}^2(p)\ \hat{\sigma}^2(f) + \hat{\sigma}^2(p)\ \hat{E}^2(f) + \hat{E}^2(p)\ \hat{\sigma}^2(f)$$

allows us tu write, carrying along both the "data" and the Arnold estimates for $\sigma^2(p)$ and $(f)$:

$$= \begin{bmatrix} .000321 \times 10^{-6} \\ .10122308 \times 10^{-6} \end{bmatrix} \left( \begin{matrix} 2.5 \times 10^{-5} \\ 17245 \times 10^{-} \end{matrix} + (41.527)^2 \right) + (.01715507)^2 \left( \begin{matrix} 2.5 \times 10^{-5} \\ 17245 \times 10^{-5} \end{matrix} \right)$$

Taking the smaller of each pair to produce a lower bound and the larger to yield an upper bound, we find a pair of estimates for $\sigma^2(Ca_f)$ as follows:

$$= \begin{matrix} .0008025 \times 10^{-11} \\ 1745.592 \times 10^{-11} \end{matrix} + \begin{matrix} .55356185 \times 10^{-6} \\ 174.55836 \times 10^{-6} \end{matrix} + \begin{matrix} .00073575 \times 10^{-} \\ 5.0752035 \times 10^{-} \end{matrix}$$

$$= \begin{matrix} .0000000008025 \times 10^{-5} \\ .001745592 \times 10^{-5} \end{matrix} + \begin{matrix} .055356185 \times 10^{-5} \\ 17.455836 \times 10^{-5} \end{matrix} + \begin{matrix} .00073575 \times 10^{-5} \\ 5.0752035 \times 10^{-5} \end{matrix}$$

$$= \begin{matrix} 0.5609194 \times 10^{-6}\ g^2 \\ 225.3279 \times 10^{-6}\ g^2 \end{matrix} \quad \begin{matrix} \sqrt{} = .0074895\ g = .74895\ mg \\ \sqrt{} = .01501093\ g = 15.01093\ mg \end{matrix}$$

Notice that the second term is the primary contributor to the sum in both cases. We return to this point at the end.

*Parameter:* $E(Ca_u)$. The table of daily means, CDR SMEAT URINE, gives this parameter in units of milliequivalents (meq); multiplying by $28\ \frac{mg}{meq}$, we find as the mean over the entire MBE

$$\hat{E}(Ca_u) = 246.58\ mg.$$

*Parameter:* $\sigma^2(Ca_u)$. We unfortunately do not have available the Ca analog to the c-duplicates table for N, whence we cannot "build up" an estimate of $\sigma^2(Ca_u)$ from that for $\sigma^2(c)$ and thence partition the total $\sigma^2(Ca_u)$. Instead we seem obliged to infer a value for $\hat{\sigma}^2(Ca_u)$ from the significant figures in the Ca column of the CDR SMEAT URINE table. As

mentioned above, this table column is in meq; the significant figures suggest a measurement to within an accuracy of $\pm$ 0.5 meq.

If we estimate $\hat{\sigma}^2(Ca_u) = 0.5/2$ meq

$= 0.25\ (28) = 7.0$ mg

we find

$$\hat{\sigma}^2(Ca_u)_{data} = 49.0 \text{ mg}^2$$

The Arnold report does, however, give a value

$$\hat{\sigma}^2(c_{Ca}) = 14.329859 \times 10^{-6}\ \left(\frac{mg}{ml}\right)^2$$

in order to use this value to compute a $\hat{\sigma}^2(Ca_u)$ for comparison to its data based value. We lack only a value for $\hat{E}(c_{Ca})$ which we obtain as the means of CDR's daily values of

$$\frac{\text{daily } Ca_u \text{ (meq)} \times 28\ \frac{mg}{meq}}{\text{daily urine volume}}$$

This turns out to be

$$\hat{E}(c_{Ca}) = 0.104585 \text{ mg/ml}.$$

Now using

$$\hat{\sigma}^2(Ca_u) = \hat{\sigma}^2(c_{Ca})\ \hat{\sigma}^2(V_{Ca}) + \hat{\sigma}^2(c_{Ca})\ \hat{E}^2(V_{Ca}) + \hat{E}^2(c_{Ca})\ \hat{\sigma}^2(V_{Ca})$$

yields

$$\hat{\sigma}^2(Ca_u)_{Arnold} = (14.3298 \times 10^{-6}\ (10) + (14.329859 \times 10^{-6})\ (2499.7)^2 + (.104585)^2\ (10)$$

$$= .000143298 + 89.53975 + .10938 = 89.649 \text{ mg}^2$$

which primarily reflects Arnold's estimate and is therefore an estimate almost independent of $\hat{\sigma}^2(Ca_u)_{data}$. We see that they are quite close, confirming our significant figures approach.

*Parameters:* $\hat{E}(NET_{Ca})$, $\hat{\sigma}^2(NET_{Ca})$. We can now compute the mean and variance of the net retention of Ca:

$$\hat{E}(NET_{Ca}) = \hat{E}(P_{Ca}) - \hat{E}(Ca_f) - \hat{E}(Ca_u) = .9572342\ g - .482358\ g - .24658\ g$$

$$= +\ .22829\ g$$

$$\hat{\sigma}^2(NET_{Ca}) = [169.0965 + \{^{.5609194}_{225.3279} + \{^{89.649}_{49.}] \times 10^{-6}\ g^2$$

Again, associating the smaller of both pairs to yield a lower bound and the larger to yield an upper bound, we find

$$\hat{\sigma}^2(NET_{Ca}) = \{^{218.657}_{484.0714}\} \times 10^{-6}\ g^2$$

whence the one-day MBE confidence interval in net retention , due to measurement error only, is

$$\hat{E}(NET_{Ca}) \pm \sqrt{2\ \hat{\sigma}(NET_{Ca})} = .22829 \pm .375\ g,$$

where $\overline{\hat{\sigma}(NET_{Ca})} = \sqrt{\frac{218.657 + 484.0714}{2}} \times 10^{-3}$

Here $2\ \overline{\hat{\sigma}}$ is 16% of $\hat{E}$. Thus, the measurements in a Ca-MBE are more accurate than those in an N-MBE, where this figure was 65%. The text showed that the corresponding result holds for the biological variability as well. Note that these two are independent results.

The following tables show the percent contribution to the total variance for the data based and the Arnold based estimates comprising $\hat{\sigma}^2(NET_{Ca})$.

Arnold Estimates

| Input | Feces | Urine |
|---|---|---|
| 65.2 | 0.22 | 34.6 |

Authors' Estimates

| Input | Feces | Urine |
|---|---|---|
| 38.1 | 50.8 | 11.1 |

These tables show that the discrepancy between our data-based estimate and the corresponding estimate by Arnold is indeed reflected strongly in what fraction of the total variance is due to the fecal measurements. But, the sensitivity of $\hat{\sigma}^2(NET_{Ca})$ to $\hat{\sigma}^2(P_{Ca})$ nevertheless suggests that the p-measurement has an importance for Ca-MBE that it does not for an N-MBE, i.e., that the fact of the feces being the primary excretory channel for Ca reflects itself in an increased importance to the fecal parameter p, not the fecal parameter f.

The dietary measurement is now a primary contributor to $\hat{\sigma}^2(NET_{Ca})$, unlike in the N-MBE, suggesting that attention to input measurement accuracy is more called for in a Ca-MBE than in an N-MBE, and our remarks as to the relative magnitudes of the several contributors to $\hat{\sigma}^2(P_{Ca})$ accordingly deserve more serious consideration in connection with Ca-MBE design than with N-MBE design.

Chart 8 summarized the estimates obtained by the N data-based calculation and by the Ca Arnold-based and data-based calculations. We would emphasize, of course, the relatively small contribution of measurement error to the total variance in metabolic balance experiments.

Chart 8. Model-Parameter Estimates that Differ for N and Ca

| Phase | Parameter | Unit | N Data | Ca(Arnold) | Ca(Data) |
|---|---|---|---|---|---|
| Input | E(p) | g | 16.056 | — | 0.9572342 |
| Input | $\sigma^2$(p) | $g^2$ | 46178 × $10^{-6}$ | — | 1690965 × $10^{-6}$ |
| Input | C.V. | % | 1.35 | — | 1.36 |
| Feces | E(p) | 1 | 0.04204058 | — | 0.01715507 |
| Feces | $\sigma^2$(p) | 1 | 0.6709 × $10^{-6}$ | 0.000321 × $10^{-6}$ | 0.10122308 × $10^{-6}$ |
| Feces | C.V. | % | 1.85 | 0.10 | 1.85 |
| Feces | K | g | 1.6182 | — | 0.712405 |
| Feces | E(f) | g | 38.4914 | — | 41.527 |
| Feces | $\sigma^2$(f) | $g^2$ | 148160 × $10^{-6}$ | 25 × $10^{-6}$ | 172450 × $10^{-6}$ |
| Feces | C.V. | % | 1.00 | 0.01 | 1.00 |
| Feces | E($\cdot_f$) | g | 1.0144 | — | 0.482358 |
| Feces | $\sigma^2$($\cdot_f$) | $g^2$ | 1162.609 × $10^{-6}$ | 0.5609194 × $10^{-6}$ | 225.3279 × $10^{-6}$ |
| Feces | C.V. | % | 3.37 | 0.15 | 3.11 |
| Urine | E(c) | g/l | 5.4096 | — | — |
| Urine | $\sigma^2$(c) | $(g/l)^2$ | 21890 × $10^{-6}$ | 14.329859 × $10^{-6}$ | — |
| Urine | C.V. | % | 2.69 | 3.17 | — |
| Urine | E($\cdot_u$) | g | 13.721 | — | 0.24658 |
| Urine | $\sigma^2$($\cdot_u$) | $g^2$ | 137100 × $10^{-6}$ | 89.649 × $10^{-6}$ | 49 × $10^{-6}$ |
| Urine | C.V. | % | 2.70 | 3.8 | 2.83 |

Note: The C.V.'s in the "Ca(Arnold)" column used the corresponding E(•) from the "Ca(Data)" column as its denominator.